

日 本 国 特 許 庁
JAPAN PATENT OFFICE

20.03.03

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出 願 年 月 日
Date of Application:

2002年 6月19日

REC'D 16 MAY 2003

WIPO PCT

出 願 番 号
Application Number:

特願2002-178169

[ST.10/C]:

[JP 2002-178169]

出 願 人
Applicant(s):

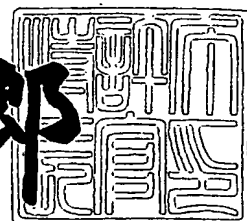
三菱電機株式会社

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH
RULE 17.1(a) OR (b)

2003年 5月 2日

特 許 庁 長 官
Commissioner,
Japan Patent Office

太田信一郎



出証番号 出証特2003-3031642

【書類名】 特許願

【整理番号】 539982JP02

【提出日】 平成14年 6月19日

【あて先】 特許庁長官 殿

【国際特許分類】 H04N 5/76

【発明者】

【住所又は居所】 東京都千代田区丸の内二丁目2番3号 三菱電機株式会社
社内

【氏名】 守屋 芳美

【発明者】

【住所又は居所】 東京都千代田区丸の内二丁目2番3号 三菱電機株式会社
社内

【氏名】 西川 博文

【発明者】

【住所又は居所】 東京都千代田区丸の内二丁目2番3号 三菱電機株式会社
社内

【氏名】 関口 俊一

【発明者】

【住所又は居所】 東京都千代田区丸の内二丁目2番3号 三菱電機株式会社
社内

【氏名】 山田 悦久

【発明者】

【住所又は居所】 東京都千代田区丸の内二丁目2番3号 三菱電機株式会社
社内

【氏名】 乙井 研二

【発明者】

【住所又は居所】 東京都千代田区丸の内二丁目2番3号 三菱電機株式会社
社内

【氏名】 黒田 慎一

【発明者】

【住所又は居所】 東京都千代田区丸の内二丁目 2 番 3 号 三菱電機株式会社内

【氏名】 小川 文伸

【発明者】

【住所又は居所】 東京都千代田区丸の内二丁目 2 番 3 号 三菱電機株式会社内

【氏名】 浅井 光太郎

【特許出願人】

【識別番号】 000006013

【氏名又は名称】 三菱電機株式会社

【代理人】

【識別番号】 100102439

【弁理士】

【氏名又は名称】 宮田 金雄

【選任した代理人】

【識別番号】 100092462

【弁理士】

【氏名又は名称】 高瀬 彌平

【先の出願に基づく優先権主張】

【出願番号】 特願2002-110259

【出願日】 平成14年 4月12日

【手数料の表示】

【予納台帳番号】 011394

【納付金額】 21,000円

【その他】

国等の委託研究の成果に係る特許出願（平成13年度、通信・放送機構、情報家電コンテンツ表現システムの研究開発委託研究、産業活力再生特別措置法第30条の適用を受けるもの）

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9704079

【ブルーフの要否】 要

【書類名】 明細書

【発明の名称】 メタデータ編集装置およびその方法、メタデータ再生装置およびその方法、メタデータ配信装置およびその方法

【特許請求の範囲】

【請求項1】 動画像と音声のいずれか一方、または双方を含むマルチメディアコンテンツを1つ以上のシーンに分割し、分割したシーン毎にメタデータを生成するメタデータ編集装置であって、

上記1つ以上のシーンから構成されるマルチメディアコンテンツの階層的な構造を記述するシーン構造情報メタデータを生成するシーン構造記述部、

を有することを特徴とするメタデータ編集装置。

【請求項2】 請求項1記載のメタデータ編集装置において、さらに、上記1つ以上に分割された各シーンの中から各シーンがもつ特徴量を抽出し、特徴記述メタデータを生成する特徴抽出部を有することを特徴とするメタデータ編集装置。

【請求項3】 請求項1または請求項2記載のメタデータ編集装置において、さらに、

各シーンの1フレーム以上の代表フレームを画像データとして抽出し、抽出した代表フレーム画像に関するメタデータを生成する代表フレーム生成部を有することを特徴とするメタデータ編集装置。

【請求項4】 請求項1～請求項3いずれかの請求項に記載のメタデータ編集装置において、さらに、

各シーンに対してテキスト情報を付与し、テキスト情報メタデータを生成するテキスト情報付与部を有することを特徴とするメタデータ編集装置。

【請求項5】 請求項1～請求項4いずれかの請求項に記載のメタデータ編集装置において、さらに、生成された各メタデータをMPEG-7に規定されたフォーマットに従い記述するメタデータ記述部を有することを特徴とするメタデータ編集装置。

【請求項6】 請求項1～請求項5いずれかの請求項に記載のメタデータ編集装置において、さらに、シーンチェンジ個所を検出するシーンチェンジ検出部

を有することを特徴とするメタデータ編集装置。

【請求項 7】 動画像と音声のいずれか一方、または双方を含むマルチメディアコンテンツを1つ以上のシーンに分割し、分割したシーン毎にメタデータを生成する際、

1つ以上のシーンから構成されるマルチメディアコンテンツの階層的な構造を記述するシーン構造情報メタデータを生成することを特徴とするコンテンツ編集方法。

【請求項 8】 動画像と音声のいずれか一方、または双方を含むマルチメディアコンテンツの構造や特徴が規定のフォーマットに従って記述されたメタデータを入力して解析するメタデータ解析手段と、

上記メタデータ解析部の解析結果を入力して、コンテンツの階層的なシーン構造を表示する構造表示部と、

を有することを特徴とするメタデータ再生装置。

【請求項 9】 請求項 8 記載のメタデータ再生装置において、さらに、

上記メタデータ解析部の解析結果によりメタデータから取得されるシーンまたはシーングループの特徴量やテキスト情報に基づいて、シーンまたはシーングループを検索する検索手段を有することを特徴とするメタデータ再生装置。

【請求項 10】 請求項 8 または請求項 9 記載のメタデータ再生装置において、さらに、

各シーンの重要度に基づいてシーン構造の要約を作成する要約作成部と、

要約作成部によって作成されたシーン構造の要約を表示する要約構造表示部とを有することを特徴とするメタデータ再生装置。

【請求項 11】 請求項 8 ～請求項 10 いずれかの請求項に記載のメタデータ再生装置において、さらに、

表示されているシーン構造の中から選択されたシーンにより構成されるマルチメディアコンテンツを再生するコンテンツ再生部を有することを特徴とするメタデータ再生装置。

【請求項 12】 動画像と音声のいずれか一方、または双方を含むマルチメディアコンテンツの構造や特徴が規定のフォーマットに従って記述されたメタデ

ータを入力して解析し、その解析結果を入力して、コンテンツの階層的なシーン構造を表示することを特徴とするメタデータ再生方法。

【請求項13】 動画像と音声のいずれか一方、または双方を含むマルチメディアコンテンツの構造や特徴が規定のフォーマットに従って記述されたメタデータを入力して解析するメタデータ解析手段と、

上記メタデータ解析手段によって解析された情報を入力して、コンテンツを配信すべき各端末の能力に基づきメタデータを再生成するメタデータ再生成部と、

上記メタデータ再生成部によって再生成されたメタデータを各端末に配信するメタデータ配信部と、

を有することを特徴とするメタデータ配信装置。

【請求項14】 請求項12記載のメタデータ配信サーバにおいて、さらに、メタデータに対応するコンテンツを配信するコンテンツ配信部を有することを特徴とするメタデータ配信装置。

【請求項15】 請求項13記載のメタデータ配信サーバにおいて、コンテンツ配信サーバは、コンテンツを配信するときに、各端末からの要求条件に応じて、コンテンツのフォーマットを変換して配信することを特徴とするメタデータ配信装置。

【請求項16】 請求項7に記載のコンテンツ編集方法により生成されたメタデータを入力して解析し、解析した情報からコンテンツを配信すべき各端末の能力に基づきメタデータを再生成して、再生成したメタデータを各端末に配信することを特徴とするメタデータ配信方法。

【請求項17】 コンテンツの意味内容、構造、特徴を記述する1つまたは複数の記述子から成るメタデータを操作する際のメタデータ最適化ヒント情報として、メタデータに含まれる記述子の名称または識別子を記述することを特徴とするメタデータ最適化ヒント情報記述方法。

【請求項18】 コンテンツの意味内容、構造、特徴を記述する1つまたは複数の記述子から成るメタデータを操作する際のメタデータ最適化ヒント情報として、メタデータに含まれる各記述子とり得る値の範囲を記述することを特徴とするメタデータ最適化ヒント情報記述方法。

【請求項19】 コンテンツの意味内容、構造、特徴を記述する1つまたは複数の記述子から成るメタデータを操作する際のメタデータ最適化ヒント情報として、メタデータに含まれる各記述子の出現個数と出現位置を記述することを特徴とするメタデータ最適化ヒント情報記述方法。

【請求項20】 請求項19記載のメタデータ最適化ヒント情報記述方法において、メタデータに含まれる各記述子の出現位置をその記述子固有の識別子またはその記述子が出現する階層構造の深さの最大値で記述することを特徴とするメタデータ最適化ヒント情報記述方法。

【請求項21】 コンテンツの意味内容、構造、特徴を記述する1つまたは複数の記述子から成るメタデータを操作する際のメタデータ最適化ヒント情報として、メタデータに含まれる記述子が階層的に構成されるときに、階層構造の深さの最大値を記述することを特徴とするメタデータ最適化ヒント情報記述方法。

【請求項22】 コンテンツの意味内容、構造、特徴を記述する1つまたは複数の記述子から成るメタデータを操作する際のメタデータ最適化ヒント情報として、メタデータに含まれる記述子の出現個数の総和を記述することを特徴とするメタデータ最適化ヒント情報記述方法。

【請求項23】 コンテンツの意味内容、構造、特徴を記述する1つまたは複数の記述子から成るメタデータを操作する際のメタデータ最適化ヒント情報として、メタデータのファイルサイズとファイルフォーマット、およびメタデータの文法を規定するシンタックスファイルの所在を記述することを特徴とするメタデータ最適化ヒント情報記述方法。

【請求項24】 コンテンツの意味内容、構造、特徴を記述する1つまたは複数の記述子から成るメタデータを操作する際のメタデータ最適化ヒント情報として、メタデータの文法を規定するシンタックスファイルの所在と、該シンタックスファイルに照らして子要素を持ちうると判断される記述子が、実際にメタデータ内ですべての子要素を含んでいるか否かを指示する情報を記述することを特徴とするメタデータ最適化ヒント情報記述方法。

【請求項25】 コンテンツの意味内容、構造、特徴を記述する1つまたは複数の記述子から成るメタデータから、メタデータ受信装置の能力や通信環境、

ユーザ要求のいずれかに適した記述のみを抽出して再構成するメタデータ再生成装置において、

メタデータに含まれる記述子の名称または識別子を記述したメタデータ最適化ヒント情報を解析するヒント情報解析手段と、

解析された上記メタデータ最適化ヒント情報を用いて、メタデータに再生成後のメタデータに必要な記述が含まれているかどうかを判断するメタデータ解析手段と、

を備えたことを特徴とするメタデータ再生成装置。

【請求項 26】 メタデータに含まれる記述子の出現位置と出現頻度を記述したメタデータ最適化ヒント情報を解析する手段と、

解析された上記メタデータ最適化ヒント情報を用いて、出現位置に対応する記述子をすべて解析したか、または出現頻度分の解析を終了したかを判断するメタデータ解析手段と、

を備えたことを特徴とする請求項 25 に記載のメタデータ再生成装置。

【請求項 27】 メタデータの再生成後に、再生成されたメタデータに対応するメタデータ最適化ヒント情報を再生成する手段を備えたことを特徴とする、請求項 25 または請求項 26 に記載のメタデータ再生成装置。

【請求項 28】 コンテンツの意味内容、構造、特徴を記述する 1 つまたは複数の記述子から成るメタデータから、メタデータ受信装置の能力や通信環境、ユーザ要求のいずれかに適した記述のみを抽出して再構成したメタデータを配信するメタデータ配信装置において、

メタデータに含まれる記述子の名称または識別子を記述したメタデータ最適化ヒント情報を解析するヒント情報解析手段と、

解析された上記メタデータ最適化ヒント情報を用いて、メタデータに再生成後のメタデータに必要な記述が含まれているかどうかを判断するメタデータ解析手段と、

を備えたことを特徴とするメタデータ配信装置。

【請求項 29】 メタデータに含まれる記述子の出現位置と出現頻度を記述したメタデータ最適化ヒント情報を解析する手段と、

解析された上記メタデータ最適化ヒント情報を用いて、出現位置に対応する記述子をすべて解析したか、または出現頻度分の解析を終了したかを判断するメタデータ解析手段と、

を備えたことを特徴とする請求項28に記載のメタデータ配信装置。

【請求項30】 コンテンツの意味内容、構造、特徴を記述する1つまたは複数の記述子から成るメタデータを用いて、コンテンツを検索するコンテンツ検索装置において、

メタデータに含まれる記述子の名称または識別子を記述したメタデータ最適化ヒント情報を解析するヒント情報解析手段と、

解析された上記メタデータ最適化ヒント情報を用いて、検索に用いる記述子が含まれているかどうかを判断するメタデータ解析手段と、

を備えたことを特徴とするコンテンツ検索装置。

【請求項31】 メタデータに含まれる記述子の出現位置と出現頻度を記述したメタデータ最適化ヒント情報を解析する手段と、

解析された上記メタデータ最適化ヒント情報を用いて、出現位置に対応する記述子をすべて解析したか、または出現個数分の解析を終了したかを判断するメタデータ解析手段と、

を備えたことを特徴とする請求項30に記載のコンテンツ検索装置。

【請求項32】 請求項17から請求項24のいずれかに記載のメタデータ最適化ヒント情報記述方法により生成されたメタデータ最適化ヒント情報と、メタデータ再生成に関わる外部条件とを取得し、取得したメタデータ最適化ヒント情報を解析する手段と、

解析された上記メタデータ最適化ヒント情報と上記メタデータ再生成に関わる外部条件とを用いて、メタデータ再生成に関わる条件を設定する手段と、

を備えたことを特徴とするメタデータ再生成条件設定装置。

【請求項33】 1つ以上のコンテンツをコンテンツ受信装置の能力や通信環境、ユーザ要求のいずれかに適したコンテンツに再生成するコンテンツ再生成装置において、

コンテンツの意味内容、構造、特徴を記述する1つまたは複数の記述子から成

るメタデータに含まれる記述子の名称または識別子を記述したメタデータ最適化ヒント情報を解析するヒント情報解析手段と、

解析された上記メタデータ最適化ヒント情報を用いて、コンテンツの再再生に必要な記述が含まれているかどうかを判断するメタデータ解析手段と、

を備えたことを特徴とするコンテンツ再生装置。

【請求項34】 コンテンツの意味内容、構造、特徴を記述する1つまたは複数の記述子から成るメタデータから、メタデータ受信装置の能力や通信環境、ユーザ要求のいずれかに適した記述のみを抽出して再構成するメタデータ再生方法において、

メタデータに含まれる記述子の名称または識別子を記述したメタデータ最適化ヒント情報を解析し、解析した上記メタデータ最適化ヒント情報を用いて、メタデータに再生後のメタデータに必要な記述が含まれているかどうかを判断することを特徴とするメタデータ再生方法。

【請求項35】 1つ以上のコンテンツをコンテンツ受信装置の能力や通信環境、ユーザ要求のいずれかに適したコンテンツに再生するコンテンツ再生方法において、

コンテンツの意味内容、構造、特徴を記述する1つまたは複数の記述子から成るメタデータに含まれる記述子の名称または識別子を記述したメタデータ最適化ヒント情報を解析し、解析した上記メタデータ最適化ヒント情報を用いて、コンテンツの再再生に必要な記述が含まれているかどうかを判断することを特徴とするコンテンツ再生方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、動画像や音声を含むマルチメディアコンテンツを複数のシーンに分割し、分割したシーン毎にメタデータを生成するメタデータ編集装置およびその方法と、そのメタデータを用いてコンテンツを再生するメタデータ再生装置およびその方法と、そのために必要なメタデータを配信するメタデータ配信サーバおよびその方法、さらにはメタデータ最適化ヒント情報記述方法、メタデータ再生

成装置、コンテンツ検索装置、メタデータ再生条件設定装置、コンテンツ再生装置、メタデータ再生方法、コンテンツ再生方法に関する。

【0002】

【従来の技術】

従来技術として、例えば、特開平2001-028722号公報に開示されている動画像管理装置がある。この動画像管理装置では、ビデオデータを複数のシーンに分割し、各シーンに対して、そのシーンの再生に必要な区間情報と、シーン番号と、シーンを代表する画像との集まりであるメタデータを作成し、各インデックスに検索目的に応じたタイトルを付与するインデックス作成手段を有する。ユーザはタイトルを用いてメタデータを検索し、インデックスに含まれるシーン番号の順番と区間情報により、ビデオデータの中の必要なシーンだけをつなぎ合わせて再生することができる。

【0003】

【発明が解決しようとする課題】

しかし、上述の従来技術では、シーンの再生に必要な区間情報と、シーン番号と、シーンを代表する画像とによりメタデータを作成するに過ぎないため、ビデオデータがもっている階層性など、ビデオデータの構造まで管理することができない、という問題があった。

【0004】

また、登録した画像を検索する際には、インデックスに付与されているタイトルを用いて行うため、適切な検索結果を得るためには、適切なタイトルを入力しなければならない、という問題もあった。

【0005】

本発明は、このような問題を解決するためになされたものであり、シーンの区間情報やタイトルのほかに、ビデオデータ等のコンテンツのもつ構造等のインデックス情報であるメタデータを生成することができるメタデータ編集装置およびその方法、さらにはメタデータ最適化ヒント情報記述方法、メタデータ再生装置、コンテンツ検索装置、メタデータ再生条件設定装置、コンテンツ再生装置、メタデータ再生方法、コンテンツ再生方法を提供することを目的とする。

【0006】

また、メタデータ編集装置およびその方法により生成されたメタデータを用いて、ユーザが見たいシーンだけを集めて再生したり、メタデータに記述されている特徴量等を用いて、見たいシーンを検索することができるメタデータ再生装置およびその方法を、さらにはそのために必要なメタデータを配信するメタデータ配信サーバおよびその方法、さらにはメタデータ最適化ヒント情報記述方法、メタデータ再生装置、コンテンツ検索装置、メタデータ再生条件設定装置、コンテンツ再生装置、メタデータ再生方法、コンテンツ再生方法を提供することを目的とする。

【0007】

【課題を解決するための手段】

上記目的を達成するため、本発明では、動画像と音声のいずれか一方、または双方を含むマルチメディアコンテンツを1つ以上のシーンに分割し、分割したシーン毎にメタデータを生成するメタデータ編集装置であって、上記1つ以上のシーンから構成されるマルチメディアコンテンツの階層的な構造を記述するシーン構造情報メタデータを生成するシーン構造記述部、を有することにより、マルチメディアコンテンツの構造を記述するインデックス情報であるメタデータを生成するメタデータ編集装置であることを特徴とする。

【0008】

特に、さらに、上記1つ以上に分割された各シーンの中から各シーンがもつ特徴量を抽出し、特徴記述メタデータを生成する特徴抽出部を有することにより、さらにマルチメディアコンテンツの特徴を記述するインデックス情報であるメタデータを生成するメタデータ編集装置であることを特徴とする。

【0009】

また、さらに、各シーンの1フレーム以上の代表フレームを画像データとして抽出し、抽出した代表フレーム画像に関するメタデータを生成する代表フレーム生成部を有することにより、コンテンツの内容把握を目的としたインデックス情報であるメタデータを生成するメタデータ編集装置であることを特徴とする。

【0010】

また、さらに、各シーンに対してテキスト情報を付与し、テキスト情報メタデータを生成するテキスト情報付与部を有することにより、検索に有用なインデックス情報であるメタデータを生成するメタデータ編集装置であることを特徴とする。

【0011】

また、さらに、生成された各メタデータをMPEG-7に規定されたフォーマットに従い記述するメタデータ記述部を有するメタデータ編集装置であることを特徴とする。

【0012】

また、さらに、シーンチェンジ個所を検出するシーンチェンジ検出部を有することにより、人手によらず自動的にシーンチェンジ個所を検出するメタデータ編集装置であることを特徴とする。

【0013】

また、動画像と音声のいずれか一方、または双方を含むマルチメディアコンテンツを1つ以上のシーンに分割し、分割したシーン毎にメタデータを生成する際、1つ以上のシーンから構成されるマルチメディアコンテンツの階層的な構造を記述するシーン構造情報メタデータを生成するコンテンツ編集方法であることを特徴とする。

【0014】

また、動画像と音声のいずれか一方、または双方を含むマルチメディアコンテンツの構造や特徴が規定のフォーマットに従って記述されたメタデータを入力して解析するメタデータ解析手段と、上記メタデータ解析部の解析結果を入力して、コンテンツの階層的なシーン構造を表示する構造表示部と、を有するメタデータ再生装置であることを特徴とする。

【0015】

また、さらに、上記メタデータ解析部の解析結果によりメタデータから取得されるシーンまたはシーングループの特徴量やテキスト情報に基づいて、シーンまたはシーングループを検索する検索手段を有するメタデータ再生装置であること

を特徴とする。

【0016】

また、さらに、各シーンの重要度に基づいてシーン構造の要約を作成する要約作成部と、要約作成部によって作成されたシーン構造の要約を表示する要約構造表示部とを有するメタデータ再生装置であることを特徴とする。

【0017】

また、表示されているシーン構造の中から選択されたシーンにより構成されるマルチメディアコンテンツを再生するコンテンツ再生部を有するメタデータ再生装置であることを特徴とする。

【0018】

また、動画像と音声のいずれか一方、または双方を含むマルチメディアコンテンツの構造や特徴が規定のフォーマットに従って記述されたメタデータを入力して解析し、その解析結果を入力して、コンテンツの階層的なシーン構造を表示するメタデータ再生方法であることを特徴とする。

【0019】

また、動画像と音声のいずれか一方、または双方を含むマルチメディアコンテンツの構造や特徴が規定のフォーマットに従って記述されたメタデータを入力して解析するメタデータ解析手段と、上記メタデータ解析手段によって解析された情報を入力して、コンテンツを配信すべき各端末の能力に基づきメタデータを再生成するメタデータ再生成部と、上記メタデータ再生成部によって再生成されたメタデータを各端末に配信するメタデータ配信部と、を有するメタデータ配信装置であることを特徴とする。

【0020】

また、さらに、メタデータに対応するコンテンツを配信するコンテンツ配信部を有するメタデータ配信装置であることを特徴とする。

【0021】

また、コンテンツ配信サーバは、コンテンツを配信するときに、各端末からの要求条件に応じて、コンテンツのフォーマットを変換して配信することを特徴とする。

【 0 0 2 2 】

また、上述のコンテンツ編集方法により生成されたメタデータを入力して解析し、解析した情報からコンテンツを配信すべき各端末の能力に基づきメタデータを再生成して、再生成したメタデータを各端末に配信するメタデータ配信方法であることを特徴とする。

【 0 0 2 3 】

また、コンテンツの意味内容、構造、特徴を記述する 1 つまたは複数の記述子から成るメタデータを操作する際のメタデータ最適化ヒント情報として、メタデータに含まれる記述子の名称または識別子を記述するメタデータ最適化ヒント情報記述方法であることを特徴とする。

【 0 0 2 4 】

また、コンテンツの意味内容、構造、特徴を記述する 1 つまたは複数の記述子から成るメタデータを操作する際のメタデータ最適化ヒント情報として、メタデータに含まれる各記述子がとり得る値の範囲を記述するメタデータ最適化ヒント情報記述方法であることを特徴とする。

【 0 0 2 5 】

また、コンテンツの意味内容、構造、特徴を記述する 1 つまたは複数の記述子から成るメタデータを操作する際のメタデータ最適化ヒント情報として、メタデータに含まれる各記述子の出現個数と出現位置を記述するメタデータ最適化ヒント情報記述方法であることを特徴とする。

【 0 0 2 6 】

特に、そのメタデータ最適化ヒント情報記述方法において、メタデータに含まれる各記述子の出現位置をその記述子固有の識別子またはその記述子が出現する階層構造の深さの最大値で記述するメタデータ最適化ヒント情報記述方法であることを特徴とする。

【 0 0 2 7 】

また、コンテンツの意味内容、構造、特徴を記述する 1 つまたは複数の記述子から成るメタデータを操作する際のメタデータ最適化ヒント情報として、メタデータに含まれる記述子が階層的に構成されるときに、階層構造の深さの最大値を

記述するメタデータ最適化ヒント情報記述方法であることを特徴とする。

【0028】

また、コンテンツの意味内容、構造、特徴を記述する1つまたは複数の記述子から成るメタデータを操作する際のメタデータ最適化ヒント情報として、メタデータに含まれる記述子の出現個数の総和を記述するメタデータ最適化ヒント情報記述方法であることを特徴とする。

【0029】

また、コンテンツの意味内容、構造、特徴を記述する1つまたは複数の記述子から成るメタデータを操作する際のメタデータ最適化ヒント情報として、メタデータのファイルサイズとファイルフォーマット、およびメタデータの文法を規定するシンタックスファイルの所在を記述するメタデータ最適化ヒント情報記述方法であることを特徴とする。

【0030】

また、コンテンツの意味内容、構造、特徴を記述する1つまたは複数の記述子から成るメタデータを操作する際のメタデータ最適化ヒント情報として、メタデータの文法を規定するシンタックスファイルの所在と、該シンタックスファイルに照らして子要素を持ちうると判断される記述子が、実際にメタデータ内ですべての子要素を含んでいるか否かを指示する情報を記述するメタデータ最適化ヒント情報記述方法であることを特徴とする。

【0031】

また、コンテンツの意味内容、構造、特徴を記述する1つまたは複数の記述子から成るメタデータから、メタデータ受信装置の能力や通信環境、ユーザ要求のいずれかに適した記述のみを抽出して再構成するメタデータ再生成装置において、メタデータに含まれる記述子の名称または識別子を記述したメタデータ最適化ヒント情報を解析するヒント情報解析手段と、解析された上記メタデータ最適化ヒント情報を用いて、メタデータに再生成後のメタデータに必要な記述が含まれているかどうかを判断するメタデータ解析手段と、を備えたメタデータ再生成装置であることを特徴とする。

【0032】

特に、そのメタデータ再生成装置において、メタデータに含まれる記述子の出現位置と出現頻度を記述したメタデータ最適化ヒント情報を解析する手段と、解析された上記メタデータ最適化ヒント情報を用いて、出現位置に対応する記述子をすべて解析したか、または出現頻度分の解析を終了したかを判断するメタデータ解析手段と、を備えたメタデータ再生成装置であることを特徴とする。

【 0 0 3 3 】

さらに、メタデータの再生成後に、再生成されたメタデータに対応するメタデータ最適化ヒント情報を再生成する手段を備えたメタデータ再生成装置であることを特徴とする。

【 0 0 3 4 】

また、コンテンツの意味内容、構造、特徴を記述する1つまたは複数の記述子から成るメタデータから、メタデータ受信装置の能力や通信環境、ユーザ要求のいずれかに適した記述のみを抽出して再構成したメタデータを配信するメタデータ配信装置において、メタデータに含まれる記述子の名称または識別子を記述したメタデータ最適化ヒント情報を解析するヒント情報解析手段と、解析された上記メタデータ最適化ヒント情報を用いて、メタデータに再生成後のメタデータに必要な記述が含まれているかどうかを判断するメタデータ解析手段と、を備えたメタデータ配信装置であることを特徴とする。

【 0 0 3 5 】

特に、そのメタデータ配信装置において、メタデータに含まれる記述子の出現位置と出現頻度を記述したメタデータ最適化ヒント情報を解析する手段と、解析された上記メタデータ最適化ヒント情報を用いて、出現位置に対応する記述子をすべて解析したか、または出現頻度分の解析を終了したかを判断するメタデータ解析手段と、を備えたメタデータ配信装置であることを特徴とする。

【 0 0 3 6 】

また、コンテンツの意味内容、構造、特徴を記述する1つまたは複数の記述子から成るメタデータを用いて、コンテンツを検索するコンテンツ検索装置において、メタデータに含まれる記述子の名称または識別子を記述したメタデータ最適化ヒント情報を解析するヒント情報解析手段と、解析された上記メタデータ最適

化ヒント情報を用いて、検索に用いる記述子が含まれているかどうかを判断するメタデータ解析手段と、を備えたコンテンツ検索装置であることを特徴とする。

【0037】

特に、そのコンテンツ検索装置において、メタデータに含まれる記述子の出現位置と出現頻度を記述したメタデータ最適化ヒント情報を解析する手段と、解析された上記メタデータ最適化ヒント情報を用いて、出現位置に対応する記述子をすべて解析したか、または出現個数分の解析を終了したかを判断するメタデータ解析手段と、を備えたコンテンツ検索装置であることを特徴とする。

【0038】

また、以上説明したいずれかのメタデータ最適化ヒント情報記述方法により生成されたメタデータ最適化ヒント情報と、メタデータ再生成に関わる外部条件とを取得し、取得したメタデータ最適化ヒント情報を解析する手段と、解析された上記メタデータ最適化ヒント情報と上記メタデータ再生成に関わる外部条件とを用いて、メタデータ再生成に関わる条件を設定する手段と、を備えたメタデータ再生成条件設定装置であることを特徴とする。

【0039】

また、1つ以上のコンテンツをコンテンツ受信装置の能力や通信環境、ユーザ要求のいずれかに適したコンテンツに再生成するコンテンツ再生成装置において、コンテンツの意味内容、構造、特徴を記述する1つまたは複数の記述子から成るメタデータに含まれる記述子の名称または識別子を記述したメタデータ最適化ヒント情報を解析するヒント情報解析手段と、解析された上記メタデータ最適化ヒント情報を用いて、コンテンツの再再生に必要な記述が含まれているかどうかを判断するメタデータ解析手段と、を備えたコンテンツ再生成装置であることを特徴とする。

【0040】

また、コンテンツの意味内容、構造、特徴を記述する1つまたは複数の記述子から成るメタデータから、メタデータ受信装置の能力や通信環境、ユーザ要求のいずれかに適した記述のみを抽出して再構成するメタデータ再生成方法において、メタデータに含まれる記述子の名称または識別子を記述したメタデータ最適化

ヒント情報を解析し、解析した上記メタデータ最適化ヒント情報を用いて、メタデータに再生成後のメタデータに必要な記述が含まれているかどうかを判断するメタデータ再生成方法であることを特徴とする。

【0041】

また、1つ以上のコンテンツをコンテンツ受信装置の能力や通信環境、ユーザ要求のいずれかに適したコンテンツに再生成するコンテンツ再生成方法において、コンテンツの意味内容、構造、特徴を記述する1つまたは複数の記述子から成るメタデータに含まれる記述子の名称または識別子を記述したメタデータ最適化ヒント情報を解析し、解析した上記メタデータ最適化ヒント情報を用いて、コンテンツの再再生に必要な記述が含まれているかどうかを判断するコンテンツ再生成方法であることを特徴とする。

【0042】

【発明の実施の形態】

実施の形態1.

本実施の形態1では、動画像や音声を含むマルチメディアコンテンツを複数のシーンに分割し、シーンの階層的な構造記述や、各シーンの特徴量を含む本発明のインデックス情報としてメタデータを作成するメタデータ編集装置について説明する。

【0043】

図1は、本実施の形態1におけるメタデータ編集装置の構成を示す。

図1において、2は編集対象のコンテンツを再生・表示するコンテンツ再生・表示部、3はコンテンツを複数のシーンに分割するシーン分割部、4はシーンの代表フレームをサムネイル画像として抽出するサムネイル画像生成部、5はシーン分割部3により分割されたシーンのグループ化、シーンの結合、シーンの削除、シーンの関係情報の生成などによりシーンを階層的に編集するシーン編集部、6は各シーンに対して各種のテキスト情報を付加するテキスト情報付与部、7はシーンの特徴を抽出する特徴抽出部、8はコンテンツ再生・表示部2、シーン分割部3、サムネイル画像生成部4、シーン記述編集部5およびテキスト情報付与部6等に対してユーザからの指示情報を入力するユーザ入力部、9はシーン分割部

3、サムネイル画像生成部4、シーン編集部5、テキスト情報付与部6および特徴抽出部7より出力される各種メタデータ情報を統合し、規定のフォーマットに従って記述したメタデータを生成するメタデータ記述部、である。また、10は映像データや音声データ等からなるマルチメディアコンテンツ、11はユーザ入力情報、12はシーンの区間情報メタデータ、13はシーンのサムネイル画像情報メタデータ、14はシーン構造情報メタデータ、15はテキスト情報メタデータ、16は特徴記述メタデータ、17はマルチメディアコンテンツの内容及び構造を記述したメタデータである。

【0044】

次に、本実施の形態1におけるメタデータ編集装置の動作を、図2に示す構成のニュース映像を編集する場合を例に説明する。

【0045】

まず、コンテンツ蓄積部（図示せず）に蓄積されている映像コンテンツ等のマルチメディアコンテンツ10がネットワーク等を介し本装置に入力すると、編集用にコンテンツ再生・表示部2にて再生・表示される。

【0046】

メタデータ編集装置のユーザがその再生映像を見ながら、シーンとして切り出す位置、すなわちシーンの開始位置と、終了位置とをユーザ入力部8より入力すると、シーン分割部3は、ユーザから入力されたシーンの開始位置と終了位置とを示すシーンの区間情報メタデータ12を生成する。

【0047】

図3に、本実施の形態1のシーン分割部3が図2に示すニュース映像から生成したシーンの区間情報メタデータ12の一例を示す。図3に示すように、シーン分割部3によって、図2に示すニュース映像コンテンツから切り出された“ニュースダイジェスト”や“国内ニュース”、“国外ニュース”等の各シーンに対して、シーンの開始位置と、終了位置との区間情報を示すシーンの区間情報メタデータ12が生成される。

【0048】

シーン記述編集部5では、ユーザからユーザ入力部8を介しシーン編集の指示

があると、シーン分割部 3 からのシーンの区間情報メタデータ 12 に基づき、シーン分割部 3 で連続的に切り出されたシーンの階層的な編集を行ない、シーン構造情報メタデータ 14 として出力する。シーンの階層的編集とは、例えば、シーンのグループ化、シーンの再分割、シーンの結合、シーンの削除がある。シーンのグループ化とは、例えば、図 2 に示すニュース映像から、例えば図 3 に示すように「国内ニュース」、「国際ニュース」、「経済ニュース」等のある特定の特徴で関連あるシーンをひとつの「ニュース」グループにまとめることをいう。また、シーンの再分割とは、1 つのシーンを複数のシーンに分割することである。シーンの結合とは、複数のシーンをまとめて 1 つのシーンとすることである。

【0049】

図 4 に、このようなシーン記述編集部 5 での編集の結果、生成された映像コンテンツの階層的な構造を記述するシーン構造情報メタデータ 14 を示す。図 4 では、シーン記述編集部 5 でシーンのグループ化や、シーンの再分割、シーンの結合等のシーンの編集により、“ニュース”というシーンが、“ニュースダイジェスト”、“ニュース”、“特集”、“スポーツ”等に編集され、さらにその“ニュース”が“国内ニュース”、“国外ニュース”、“経済ニュース”等の階層的に編集されたことを示している。

【0050】

そして、シーン記述編集部 5 で生成された例えば図 4 等を示すメタデータは、メタデータ記述部 9 へ出力される。

【0051】

一方、サムネイル画像生成部 4 では、シーン分割部 3 からのシーンの区間情報メタデータ 12 に基づき、シーン分割部 3 で切り出された各シーンの中から代表フレームをサムネイル画像として生成し、生成したサムネイル情報をサムネイル画像情報メタデータ 13 としてメタデータ記述部 9 へ出力し、メタデータ記述部 9 に登録する。ここで、ユーザはユーザ入力部 8 より、サムネイルの選択を行うこともできるが、自動的に、先頭フレームや固定時間間隔の複数フレームを代表フレームとしたり、シーンチェンジ点を自動検出し、それらのフレームを代表フレームとすることもできる。サムネイル画像情報メタデータ 13 は、映像コンテンツの

中のサムネイルの位置情報(フレーム番号または時間)、または、サムネイル画像のURLなどのロケーション情報になる。

【 0 0 5 2 】

また、特徴抽出部 7 では、シーン分割部 3 からのシーンの区間情報メタデータ 12 に基づき、各シーンの中から、動き、色、あるいはシーン内に含まれるオブジェクトの形状など、シーンがもつ視覚的特徴量を抽出する。抽出された特徴量は、特徴記述メタデータ 16 として、メタデータ記述部 9 へ出力され、登録される。

【 0 0 5 3 】

また、テキスト情報付与部 6 では、シーン分割部 3 からのシーンの区間情報メタデータ 12 に基づき、各シーンに対して、ユーザがタイトル、アブストラクト、キーワード、コメント、シーンの重要度などの各種のテキスト情報を付与する。テキスト情報の付与は、ユーザ入力部 8 よりユーザが入力する場合と、コンテンツに含まれる音声情報やキャプションを解析することにより、自動的に付与する場合がある。テキスト情報は、テキスト情報メタデータ 15 として、メタデータ記述部 9 へ出力し、メタデータ記述部 9 に登録する。

【 0 0 5 4 】

図 5 に、コンテンツ再生・表示部 2 と、ユーザ入力部 8 等の画面イメージ例を示す。図 5 において、ビデオ再生画面 G 1 は、コンテンツ再生・表示部 2 における画面イメージ例に相当し、このビデオ再生画面 G 1 では、編集用のコンテンツを再生・表示する。図 5 には、明示していないが、「再生」、「停止」、「巻き戻し」、「早送り」、「コマ送り」再生用の指示ボタンなど、通常のビデオ再生装置に備えられているユーザインタフェースを備えているものとする。そして、このビデオ再生画面 G 1 の下方にはシーン分割指示画面 G 2 が表示されており、シーン分割指示画面 G 2 は、例えば、スライダー形式になっており、ビデオ再生画面 G 1 に表示される映像を見ながら、ユーザがビデオ再生画面 G 1 に表示される映像シーンの開始位置と終了位置とを指示することができる。また、シーン分割指示画面 G 2 は、同時に、シーンの開始位置と終了位置との間で、サムネイルの位置を指示することができる。ここで、シーン分割指示画面 G 2 によりサムネイルの位置が指定されると、映像コンテンツの指定された位置のフレームからサムネイル画

像生成部 4 がサムネイル画像を生成することになる。

【 0 0 5 5 】

また、シーン分割指示画面 G 2 により位置が指定されサムネイル画像は、シーンの分割情報として、シーン分割情報表示画面 G 3 に表示される。シーン分割情報表示画面 G 3 では、サムネイル画像のほか、図 3 に示すようにシーンごとにシーンの開始位置と終了位置とを示す情報を表示することもできる。

【 0 0 5 6 】

次に、ツリー構造生成指示・表示画面 G 4 では、シーンの編集をユーザに対し指示する。ユーザは、シーン分割情報表示画面 G 3 に表示されているサムネイル画像等のシーン分割情報を見ながら、映像コンテンツのもつ階層構造を表すツリーを生成する。操作方法としては、例えば、シーンのグループ化を行う場合には、ツリー上に新しいノードを追加し、そのノードにグループ化したいシーンを追加していく。シーンの追加の操作は、追加したいシーンをシーン分割情報表示画面 G 3 上で選択し、ドラッグ・アンド・ドロップにより、ノードにシーンを追加するなどの方法が考えられる。シーン分割情報表示画面 G 3、ツリー構造生成指示・表示画面 G 4 では、シーンを選択することにより、そのシーンに対してテキスト情報付与部 6 を介してテキスト情報を付与するためのユーザインタフェースがユーザ入力部 8 として提供され、シーンに対するテキスト情報を入力することができる。

【 0 0 5 7 】

メタデータ記述部 9 は、シーン分割部 3、サムネイル画像生成部 4、シーン記述編集部 5、テキスト情報付与部 6 および特徴抽出部 7 より出力される各種メタデータを統合し、規定の記述フォーマットに従って記述したメタデータファイルを生成する。メタデータの規定の記述フォーマットは、独自に定めたフォーマットで記述することもできるが、本実施の形態 1 では、ISO で標準化されている MPEG-7 を使うことにする。MPEG-7 は、コンテンツの構造や特徴を記述するフォーマットを規定するものであり、XML ファイルフォーマットとバイナリフォーマットがある。

【 0 0 5 8 】

従って、本実施の形態1のメタデータ編集装置によれば、シーンを階層的に編集するシーン記述編集部5や、シーンから特徴を抽出する特徴抽出部7を設けるようにしたので、ビデオデータ等のコンテンツのもつ階層的な構造や各シーンの特徴量を記述したメタデータを生成することができる。

【0059】

なお、コンテンツ再生・表示部2に入力されるマルチメディアコンテンツは、ネットワーク上にあるコンテンツサーバ（図示せず）から取得する場合や、メタデータ編集装置内にあるコンテンツ蓄積部（図示せず）より取得する場合、CDやDVD等の蓄積媒体（図示せず）より取得する場合など様々な場合が想定される。同様に、メタデータ記述部9より出力されるメタデータは、ネットワーク上にあるメタデータサーバ（図示せず）に蓄積される場合や、メタデータ編集装置内にあるメタデータ蓄積部（図示せず）に蓄積される場合、CDやDVD等の蓄積媒体（図示せず）にコンテンツとともに蓄積される場合などが想定される。

【0060】

また、本実施の形態1では、シーン記述編集部5と、特徴抽出部7との両者を設けて説明したが、これに限らず、シーン記述編集部5のみを設けたり、特徴抽出部7のみを設けるようにしても勿論良い。

【0061】

実施の形態2.

本実施の形態2では、実施の形態1のメタデータ編集装置により生成されたメタデータを利用して、画像の要約再生や、検索などを行うメタデータ再生装置について説明する。

【0062】

図6は、本実施の形態2におけるメタデータ再生装置の構成を示す。

図6において、19はメタデータの解析を行うメタデータ解析部、20はメタデータに記述されているコンテンツの階層的な構造を表示する構造表示部、21はサムネイル画像表示部、22は検索や再生等の指示を行うユーザ入力部、23はユーザからの検索指示に基づき検索を行う検索部、24は検索結果を表示する検索結果表示部、25はユーザからの要約作成指示に基づき要約作成を行う要約作成部、26は要

約構造表示部、27はコンテンツを再生・表示するコンテンツ再生部、である。また、28はコンテンツのもつ階層的なシーン構造や各シーンのサムネイルに関する情報、各シーンの特徴量などが記述されたメタデータ、29はメタデータ解析結果から取得されるシーン構造、30はメタデータ解析結果から取得されるサムネイル画像情報、31は検索条件、32はメタデータから取得されるシーンの特徴量やテキスト情報、33は検索結果、34は要約作成条件、35は要約情報、36はコンテンツ再生指示、37は再生するコンテンツ、38は要約されたコンテンツの構造である。

【0063】

次に動作を説明する。

まず、メタデータ解析部19は、コンテンツのもつ階層的なシーン構造や、各シーンのサムネイルに関する情報、各シーンの特徴量などが記述されたメタデータを入力とし、メタデータの解析を行う。本実施の形態2では、メタデータは実施の形態1のメタデータ記述部9が生成したMPEG-7規定のフォーマットで記述されているため、メタデータはXMLで記述されたテキストファイルでもよいし、バイナリフォーマットで符号化されたバイナリファイルであってもよい。このため、メタデータ解析部19は、メタデータがXMLで記述されているならば、XMLファイルの解析を行うXMLパーザの機能をもつ。また、メタデータがバイナリフォーマットで符号化されているならば、メタデータの復号を行うデコーダの機能を有する。

【0064】

構造表示部20は、メタデータ解析部19の解析結果を入力して、コンテンツの階層的なシーン構造を表示する。コンテンツのシーン構造は、例えば、図4に示すように、各シーンのタイトルとともにツリー表示される。

【0065】

サムネイル画像表示部21は、メタデータ解析部19の解析結果を入力して、コンテンツのサムネイル画像一覧を表示する。

【0066】

検索部23は、ユーザ入力部22を介したユーザからの検索指示により、コンテンツに含まれるシーンの検索を行う。その際、ユーザ入力部22では、キーワ

ードやサンプル画像の提示等により、検索条件を入力する。検索部 2 3 では、メタデータに記述されているシーンの特徴量や、シーンのタイトルなどのテキスト情報に基づき、ユーザから提示された条件(キーワードやサンプル画像の特徴)に合致するシーンの検索を行う。

【0067】

検索部 2 3 による検索が終了すると、検索結果表示部 2 4 は、検索部 2 3 の検索結果を入力とし、検索結果の表示を行う。検索結果の表示方法としては、例えば、検索条件に合致したシーンのサムネイル画像を表示する。

【0068】

また、要約作成部 2 5 は、ユーザ入力部 2 2 を介したユーザからの要約作成指示に基づき、コンテンツの要約を作成する。その際、ユーザ入力部 8 では、要約されたコンテンツの再生時間や、ユーザ嗜好などの情報を入力する。例えば、コンテンツがニュース映像の場合には、ニュースの中のスポーツを中心に見たい、あるいは、1時間のニュースを20分に要約してみたいなどのユーザの嗜好情報を入力する。また、要約作成部 2 5 では、メタデータに記述されているシーンの再生時間や、シーンのタイトルなどのテキスト情報に基づき、要約条件に合致した要約情報を作成する。要約情報とは、例えば、要約されたコンテンツに含まれるシーンの再生リストであり、コンテンツのURLなどのロケーション情報と、そのコンテンツの中の再生したいシーンの開始位置と終了位置を列記したリストである。

【0069】

また、コンテンツ再生・表示部 2 7 では、要約情報に含まれるコンテンツのロケーション情報に基づき、対象コンテンツを特定し、要約情報に含まれるシーンリストに基づき、再生するシーンの取得・再生・表示を行う。また、別の例として、要約情報が要約されたシーンの構造を階層的に記述したものである場合も考えられる。

【0070】

図 7 に、要約されたシーン構造を階層的に記述したものの一例を示す。

図 7 (a) は、オリジナルコンテンツのシーン構造の例を示しており、各シーン

に対して、重要度が0.0～1.0の範囲で付加されている。1.0は重要度が最も高く、0.0は重要度が最も低いことを意味する。重要度は、例えば、ユーザ嗜好に基づいて計算されたものであるとする。例えば、チームAのサッカーの試合に関するシーン、特に、試合結果と得点シーンは必ず見たいというユーザ嗜好を予め登録しておく、各シーンにはユーザ嗜好を反映した重要度が付加される。そして、図7(a)において、重要度が最も高いシーンのみで要約を生成すると、要約されたシーン構造は図7(b)に示すようになる。なお、各シーンは、そのシーンを含むコンテンツのURLなどのロケーション情報と、そのシーンのコンテンツ内での位置情報(開始位置と終了位置)などのメタデータをもつ。要約されたシーン構造に関する情報は、要約構造表示部26へ渡され、要約構造表示部26は、要約されたシーン構造を例えば、図7(b)に示すようなツリー形式で表示する。

【0071】

また、ユーザ入力部22を介しユーザが構造表示部20または要約構造表示部26に表示されているシーン構造や、サムネイル画像表示部21や検索結果表示部24に表示されている各シーンのサムネイルの中から再生したい1つ以上のシーンを選択すると、コンテンツ再生・表示部27ではコンテンツに含まれるシーンを再生・表示することができる。

【0072】

従って、本実施の形態2のメタデータ再生装置によれば、実施の形態1で説明したメタデータ編集装置により生成されたメタデータを用いて、ユーザが見たいシーンだけを集めて再生したり、メタデータに記述されている特徴量を用いて、見たいシーンを検索することができる。

また、本実施の形態1では、コンテンツ再生・表示部が、メタデータ再生装置内にあったが、コンテンツ再生・表示部が別の装置内にあってもよい。これは、例えば、シーンの構造やサムネイル画像の表示等、メタデータ再生に関わる操作や表示は、携帯電話や携帯情報端末等で行い、マルチメディアコンテンツの再生に関わる処理や表示は、携帯電話や、携帯情報端末等にネットワークを介して接続した端末(例えば、PC等)で行う場合が考えられる。

【0073】

実施の形態 3

実施の形態1では、シーンの分割をすべて手動で行っていたが、ここではシーンチェンジ点を自動検出するシーンチェンジ検出部を備えたことを特徴とする実施の形態3のメタデータ編集装置について述べる。

【0074】

図8に、本実施の形態3におけるメタデータ編集装置の構成を示す。

図8において、39はシーンチェンジ検出部、40は自動検出されたシーン開始点情報である。なお、シーンチェンジ検出部39と、シーン分割部3以外の動作は実施の形態1と同じである。

【0075】

次に動作を説明する。なお、実施の形態1のメタデータ編集装置と動作が重複するので、実施の形態2特有の動作を説明する。

【0076】

シーンチェンジ検出部39は、自動的にシーンチェンジ、カット点検出を行う。シーンチェンジ検出は、例えば、フレーム間画素差分や、フレーム間の色や、輝度のヒストグラム差分などに基づいて行う。シーン分割部3は、シーンチェンジ検出部39で検出されたシーンチェンジ点に基づき、シーンの開始位置と終了位置を決定する。

【0077】

ここで、シーンチェンジ検出部39とシーン分割部3の処理について、編集対象のコンテンツがニュース映像の場合を例に詳述する。

【0078】

シーンチェンジ検出のための特徴量として、色のヒストグラムを用いる場合を例に説明する。

【0079】

シーンチェンジ抽出部39では、1フレームごとに色のヒストグラムを算出する。表色系としては、HSV、RGB、YCbCrなどがあるが、ここでは、HSV色空間を用いる。HSV色空間は、色相(H)、彩度(S)、明度(V)の3要素から構成される。各要素のヒストグラムを算出する。次に求められたヒストグラムから、例えば下記の

式1に基づき、フレーム間のヒストグラム差分を算出する。シーンの開始フレームから、Nフレーム(例えばN=3)は、同一シーン、すなわち、シーンチェンジ点がないものと仮定する。なお、シーンの初期特徴量として、下記の式2に基づき、最初のNフレーム間のヒストグラム差分の平均(mean)と標準偏差(sd)を求める。

【0080】

【数1】

$$sum_i = \sum_{k=1}^{bin_H} |H_i(k) - H_{i-1}(k)| + \sum_{k=1}^{bin_S} |S_i(k) - S_{i-1}(k)| + \sum_{k=1}^{bin_V} |V_i(k) - V_{i-1}(k)|$$

sum_i : フレーム*i*とフレーム*i-1*のヒストグラム間差分和
 $H_i(h)$: 色相のヒストグラム, bin_H : ヒストグラムの要素数
 $S_i(h)$: 彩度のヒストグラム, bin_S : ヒストグラムの要素数
 $V_i(h)$: 明度のヒストグラム, bin_V : ヒストグラムの要素数

(式1)

【0081】

【数2】

$$mean = \frac{1}{N-1} \sum_{i=1}^{N-1} sum_i, sd = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N-1} (sum_i - mean)^2}$$

$mean$: フレーム間のヒストグラム差分の平均
 sd : フレーム間のヒストグラム差分の標準偏差

(式2)

【0082】

そして、N+1フレーム以降は、ヒストグラムのフレーム間差分が $mean + \lambda \cdot sd$ より大きくなったフレームをシーンチェンジ点とし、新しいシーンの開始点候補とする。

【0083】

シーンの開始点候補が複数得られたら、次に、ニュース映像のように、ニュースの切り替わりなどで決まったパターンの画像が挿入される場合を考える。

【0084】

ニュース映像は、ニュースの切り替わりで、例えば、アナウンサーと背景のスタジオセット、説明の文字(キャプション)により構成された画像など、決まった

パターンの画像が挿入されることが多い。従ってこれらの決まったパターンの画像(テンプレート画像と称する)、またはテンプレート画像の特徴量を記述したメタデータを予め登録しておく。テンプレート画像の特徴量とはテンプレート画像の色のヒストグラム、あるいは、動きのパターン(ニュースの切り替わりで、アナウンサーが写る部分は動きが少ないなど)などが挙げられる。テンプレート画像が予め登録されている場合には、例えば図9に示すようにシーンチェンジ点に対応する画像と、テンプレート画像とのマッチングをとり、類似度が高い場合には、そのシーンチェンジ点をシーンの開始点として登録する。類似度のマッチングとしては、フレーム間差分やフレーム間の色のヒストグラム差分などがある。また、テンプレート画像の特徴量が予め登録されている場合には、シーンチェンジ点に対応する画像から特徴量を抽出し、テンプレート画像の特徴量とのマッチングをとり、類似度が高い場合には、そのシーンチェンジ点をシーンの開始点として登録する。シーン開始点の情報はシーン分割部3へ出力される。

【0085】

シーン分割部3では、シーンチェンジ検出部39で自動抽出されたシーン開始点情報に基づき、シーンの開始点および終了点を決定する。なお、本実施の形態3のシーン分割部3では、実施の形態1と同様に、ユーザからの指示に基づき、シーンの開始点および終了点を決定することもできる。

シーン分割部3では、シーンの開始点および終了点を記述したシーンの区間情報メタデータをシーンチェンジ検出部へ出力し、シーンチェンジ検出部39にて、そのシーンに含まれるシーンチェンジ点を検出することもできる。

シーン記述編集部5では、シーン分割部3からのシーン情報に基づき、シーンチェンジ検出部で自動抽出されたシーンを再分割や統合することができる。なお、シーン記述編集部5の詳細は、実施の形態1と同様である。

【0086】

従って、本実施の形態3のメタデータ編集装置によれば、実施の形態1と同様にビデオデータ等のコンテンツのもつ階層的な構造や各シーンの特徴量を記述したメタデータを生成することができると共に、シーンチェンジ検出部39を設けたことにより、コンテンツのシーンチェンジ点を自動検出することが可能となる

【0087】

実施の形態4.

本実施の形態4では、コンテンツをクライアント端末の端末能力に応じてスケラブルに構成して配信するコンテンツ配信サーバと、そのコンテンツのメタデータをクライアント端末に対し配信するメタデータ配信サーバについて述べる。

【0088】

図10は、本実施の形態4におけるコンテンツ配信システムの構成を示す。

図10において、41は例えば実施の形態1, 3のメタデータ編集装置が生成したメタデータを蓄積したメタデータ蓄積部、42はメタデータの解析を行うメタデータ解析部、43はクライアント端末から送信される端末情報に基づいて、クライアント端末の端末能力を判定する端末能力判定部、44はクライアント端末の端末能力に応じてコンテンツを再構成し、その内容を記述したメタデータを再生成するメタデータ再生成部、45はメタデータ再生成部44で再生成されたメタデータを各種クライアント端末に配信するメタデータ配信部であり、これらにより本実施の形態4のメタデータ配信サーバを構成している。なお、メタデータ蓄積部41は、本実施の形態4のメタデータ配信サーバ外に設けられていても良く、本実施の形態4のメタデータ配信サーバはネットワーク（図示せず）等を介しメタデータ蓄積部41からメタデータを入力するように構成しても勿論良い。

【0089】

また、図10において、46はコンテンツ蓄積部、47はクライアント端末からの要求に応じてコンテンツを配信するコンテンツ配信部であり、これらによりコンテンツ配信サーバを構成する。メタデータ配信サーバの場合と同様に、コンテンツ蓄積部46はコンテンツ配信サーバ外に設けられていて、本実施の形態4のコンテンツ配信サーバはネットワーク（図示せず）を介してコンテンツデータを入力するようにしても勿論良い。

【0090】

また、図10において、481～48nは各種のクライアント端末、49はコンテンツの構造や特徴を記述したメタデータ、50はメタデータ解析結果、51はクライアン

ト端末の性能に関する情報、52は再構成されたメタデータ、53は各クライアント端末に配信されるメタデータ、54はコンテンツ配信要求、55はコンテンツ、56は各クライアント端末に配信されるコンテンツである。

【0091】

次に動作を説明する。

まず、コンテンツ配信サーバ側では、メタデータ解析部42がメタデータ蓄積部41に蓄積されているメタデータの解析を行う。メタデータ解析部42の動作は、図6に示す実施の形態2のメタデータ再生装置のメタデータ解析部19と同様で、メタデータ解析部42でメタデータを解析することにより、各コンテンツの構造や特徴に関する情報を取得する。

【0092】

図11に、ニュース映像を例に、メタデータ解析部42より出力されるコンテンツの構造情報を図示する。図11では、コンテンツの階層的なシーン構造をツリーを用いて表示している。ツリーの各ノードは、各々のシーンに対応し、各ノードには各種のシーン情報が対応付けられている。シーン情報とは、シーンのタイトル、アブストラクト、シーンの開始位置や終了位置等の時間情報等、シーンのサムネイルや代表フレームあるいはサムネイルショットや代表ショット、色や動きなど視覚的な特徴量等であるシーンの特徴などである。なお、図11には、各種のシーン情報のうち、シーンのタイトルのみを図示している。

【0093】

ここで、クライアント端末48は、端末能力の異なる各種情報家電機器を想定する。端末能力とは、通信速度、処理速度や、再生・表示可能な画像フォーマット、画像解像度、ユーザ入力機能、などである。例えば、クライアント端末481は、通信速度、処理速度、表示性能、ユーザ入力機能において、十分な性能を有するPCであり、クライアント端末482は、携帯電話、その他のクライアント端末はPDA等を想定する。各クライアント端末481～48n1～48nからは、それぞれの端末性能に関する情報が送信される。

【0094】

端末能力判定部44は、各クライアント端末481～48n1～48nから送信される端末

性能に関する情報を解析して、配信可能な画像フォーマット、最大画像解像度、コンテンツの長さ等を決定し、メタデータ再生部43へ出力する。例えば、オリジナルのコンテンツがMPEG-2で符号化された解像度の大きなビデオコンテンツの場合、十分な性能を有するクライアント端末481では、オリジナルのコンテンツを再生でき、かつ、実施の形態2で述べた画像の要約再生や検索が可能な機能を有するものとする。クライアント端末482では、MPEG-4で符号化された短いビデオショットのみ再生可能で、かつ、表示可能な最大解像度も小さいものとする。

【0095】

メタデータ再生部43では、端末能力判定部44からの各クライアント端末481～48nの端末性能に応じてコンテンツを再構成し、その構造や内容を記述したメタデータを再生し、メタデータ配信部45へ出力する。例えば、クライアント端末481に対しては、オリジナルのメタデータをそのまま配信するため、コンテンツの再構成は行わない。クライアント端末482に対しては、短いビデオショットの再生機能のみを有するため、すべてのシーンを再生することはできない。従って重要なシーンの短いビデオショットでコンテンツを再構成する。

【0096】

図12に、再構成後のコンテンツの構造例を示す。各ニュースのシーンの中から、重要なシーンを抽出し、そのシーンの代表ショットまたは、代表フレームのみで構成されるようにする。また、クライアント端末482は、実施の形態2で述べた検索機能を有しないため、シーンの各種情報のうち、検索に使用するシーンの特徴量をメタデータに記述する必要はない。そこで、メタデータ再生部43は、再構成されたシーン構造と、そのシーンの代表ショットまたは代表フレームの位置情報のみを記述したメタデータを再生し、メタデータ配信部45へ出力する。

【0097】

メタデータ配信部45は、メタデータ記述部9で生成されたメタデータをクライアント端末48に配信する。

【0098】

各クライアント端末481～48nは、メタデータ配信部45より配信されたメタデータを解析し、コンテンツのシーン構造情報を取得する。各クライアント端末481

～48nのユーザは、再生したいシーンを選択すると、選択されたシーンの位置情報が各クライアント端末481～48nからコンテンツ配信サーバのコンテンツ配信部47へ送信される。

【0099】

コンテンツ配信サーバのコンテンツ配信部47では、各クライアント端末481～48nより送信されるシーンの位置情報を取得し、コンテンツ蓄積部46より対応するコンテンツを取得し、クライアント端末48へ配信する。クライアント端末1の場合には、シーンの開始位置と終了位置を送信し、オリジナルのコンテンツの対応するシーンを配信する。クライアント端末2では、シーンの代表ショットの所在情報(URIなど)を送信する。なお、代表ショットが、クライアント端末で再生・表示不可能な画像フォーマットや、画像解像度、画像ファイルサイズ等の場合には、コンテンツ配信部47で、フォーマット変換や解像度変換、ファイルサイズを小さくするためのコンテンツの要約等を行い、送信するようにする。

【0100】

従って、本実施の形態4のメタデータ配信サーバによれば、各クライアント端末481～48nの能力に応じてメタデータを再生成して各クライアント端末に配信することが可能となる。

【0101】

なお、図10では、コンテンツ配信サーバと、メタデータ配信サーバとを別々に構成して示しているが、本発明では、これに限らず、コンテンツ配信サーバの中にメタデータ配信サーバを設けたり、あるいはメタデータ配信サーバの中にコンテンツ配信サーバを設ける等して、コンテンツ配信サーバとメタデータ配信サーバとを同一サーバ内に設けるようにしても勿論良い。このようにすれば、コンテンツ配信部47は、端末能力判定部44から各クライアント端末481～48nの能力を簡単に知ることができるので、各クライアント端末481～48nの能力に応じてコンテンツをフォーマット変換などの再構成をして各クライアント端末に配信することが可能となる。

【0102】

なお、上記実施の形態4では、メタデータ蓄積部41に蓄積されているメタデー

タは、例えば実施の形態1, 3のメタデータ編集装置が生成したメタデータとして説明したが、これに限らず、実施の形態1, 3のメタデータ編集装置以外が生成したメタデータを蓄積したものでも勿論良い。

【0103】

実施の形態5.

本実施の形態5では、実施の形態4で述べたメタデータ配信サーバの別の実施の形態を述べる。実施の形態4では、クライアント端末から送信される端末情報に基づいて、メタデータの再生成を行っていたが、本実施の形態5では、より適切なメタデータの再生成を行うために、メタデータ再生成のためのヒント情報であるメタデータ最適化ヒント情報を用いて、メタデータの再生成を行う手段を備えたことを特徴とするメタデータ配信サーバについて述べる。

【0104】

図13に、本実施の形態5におけるメタデータ配信サーバの構成を示す。49はコンテンツの構造や特徴を記述したメタデータ、53はクライアント端末に配信されるメタデータ、60はメタデータ最適化ヒント情報、61はヒント情報解析部、62は解析されたメタデータ最適化ヒント情報、63はメタデータ解析・再生成部、64は再構成されたメタデータ、65はクライアント端末の性能に関する情報、あるいは、ユーザ嗜好等のメタデータ再生成に関わる条件である。

【0105】

メタデータ蓄積部41には、コンテンツの構造や特徴を記述したメタデータ49と、そのメタデータ49を再生成するためのヒント情報のメタデータ最適化ヒント情報60が蓄積されている。メタデータ49を再生成するためのメタデータ最適化ヒント情報60とは、そのメタデータ49に、どのような種類の情報が、どのくらい含まれているか、メタデータ49の概要や複雑度を記したものである。

【0106】

図14に示す構造をもったビデオコンテンツを例に、メタデータ最適化ヒント情報60について詳述する。

【0107】

ビデオコンテンツ(Root)(Soccer game program)は、大きく2つのシーン(Scene

1, Scene 2)の前半戦と後半戦から構成され、前半戦と後半戦の各シーンは、さらに複数のシーン(Scene 1-1, Scene 1-2, Scene 1-3,)(ゴールシーン、コーナーキックシーンなど)から構成されている。図14では、シーン間の時間的な階層構造を木構造により示している。

【0108】

対応するメタデータ49には、このようなコンテンツの時間的な階層構造、すなわち、シーン間の時間的關係や各シーンの開始時間と長さ、が記述されている。また、各シーンに対しては、階層のレベルに応じてそのシーンがもつ特徴(例えば、色のヒストグラムや動きの複雑性)のほか、タイトル、アブストラクト、ジャンル、注釈等のテキスト情報、重要度、などが記述されている。なお、本実施の形態5では、メタデータの記述フォーマットとして、ISOで標準化されているMPEG-7を用いるものとする。

【0109】

図15に、MPEG-7を用いた場合のメタデータの記述例を示す。

MPEG-7では、各シーンは「ビデオセグメント」と呼ばれる単位で記述される。各ビデオセグメントには、時間情報(シーンの開始点と長さ)、タイトル、概要、ジャンル、などが記述される。なお、ビデオセグメントの階層に応じてビデオセグメント内で記述される情報が異なる場合がある。図15の例では、レベル2と3のビデオセグメントには、重要度が記述されているが、レベル4では重要度は記述されない。また色や動きの特徴量はレベル4のビデオセグメント内でのみ記述されている。

【0110】

シーン間の時間的な階層関係は、ビデオセグメントを再帰的に記述することにより表現することができる。図15の記述例では、「時間分割」という記述により、1つのビデオセグメントが時間的に分割された複数のビデオセグメントから構成されていることを記述している。MPEG-7では、コンテンツのもつ空間的な階層構造も同様に記述することができる。この場合には、「時間分割」という記述の代わりに、1つのセグメントが空間的に分割された複数のセグメントから構成されていることを示す「空間分割」という記述を用いる。

【0111】

メタデータ49の再生のためのメタデータ最適化ヒント情報60と、そのメタデータ49に含まれる情報(記述子)の種類や内容を記述したものである。図15のメタデータに対しては、コンテンツのもつ時間的な階層構造を表現する記述子(「時間分割」と、色のヒストグラムや動きの複雑性を表現する記述子、タイトル、アブストラクト、ジャンル、重要度を表現する記述子が含まれていること、また、記述の内容や複雑性を表す指標として、ビデオセグメントの階層構造の深さは最大で4(レベル1～レベル4)であること、重要度は5つの離散的な値($\{0.0, 0.25, 0.5, 0.75, 1.0\}$)をとること、重要度の視点として、“Team A”の視点で観たときの重要度と、“Team B”の視点で観たときの重要度が記述されていること、また、重要度が記述されている階層位置(ビデオセグメントのどのレベルに記述されているか)などがヒント情報になる。

【0112】

図16に、メタデータ最適化ヒント情報60のフォーマット例を示す。図16に示すメタデータ最適化ヒント情報60には、メタデータファイル情報と、メタデータ構成要素情報とが含まれる。メタデータファイル情報は、メタデータファイルの所在や、ファイルサイズ、メタデータファイルフォーマット(XMLフォーマット、バイナリフォーマットなどファイルフォーマットを示す)、シンタックスファイル情報、メタデータに含まれる要素数である出現要素数など、メタデータを処理するために必要なリソース(メタデータの蓄積・解析を行うために必要とするメモリサイズやメタデータを解析するために必要な処理系(S/W)など)を予測するための情報を記述する。なお、メタデータファイルのフォーマットを規定したシンタックスファイルとは、例えば、メタデータがXMLで記述されている場合には、その記述フォーマット(文法)を規定するDTDファイルやschemaファイルなどがそれに相当し、シンタックスファイル情報とは、DTDファイルやschemaファイルなどのシンタックスファイルの所在を記述する。

【0113】

メタデータ構成要素情報とは、メタデータを構成する記述子の種類とその内容を記述する情報である。メタデータ構成要素情報には、メタデータに含まれてい

る記述子の名前、その記述子がメタデータ内に出現する回数、その記述子が文法的に含む可能性のあるすべての記述子を含んでいるかどうかの記述(完全記述性)、またその記述子が再帰的に記述される場合には、その記述子をもつ時間または空間における階層性(深さの最大値)が含まれる。例えば、図15に示すメタデータ記述例では、「ビデオセグメント」が再帰的に記述される記述子であり、最大で4階層の構造をもっているため、「ビデオセグメント」記述子をもつ階層の深さは最大4になる。

【0114】

さらに、再帰的に記述される記述子内に含まれる記述子については、その記述子が登場する階層(出現位置)もヒント情報である。例えば、「重要度」は「ビデオセグメント」内に含まれる記述子であるが、レベル3以上の「ビデオセグメント」に含まれる、すなわち、レベル4のビデオセグメントには含まれないときに、「重要度」が登場する位置は、最大で3までになる。なお、このように、出現位置を階層レベルで指定することもできるが、「重要度」を含む「ビデオセグメント」、あるいは「ビデオセグメント」自体にIDが振られている場合には、出現位置をIDのリストとして記述することも可能である。また、値をもつ記述子の場合には記述子の型や記述子がとり得る値の範囲もヒント情報の一つである。例えば、重要度が"Team A", "Team B"のそれぞれの視点から、5つの離散的な値({0.0, 0.25, 0.5, 0.75, 1.0})で表現されている場合には、「重要度」のとり得る値は、浮動小数点の型をもつリスト{0.0, 0.25, 0.5, 0.75, 1.0}である。以上の記述をメタデータの構成要素である記述子毎に繰り返す。

【0115】

図17に、図16のフォーマットに従って記述されたメタデータ最適化ヒント情報の一例を示す。図17に示すメタデータ最適化ヒント情報60の一例には、メタデータファイル情報と、「ビデオセグメント」と「タイトル」というメタデータ構成要素情報とが含まれていることがわかる。

【0116】

次にメタデータ最適化ヒント情報を使って、メタデータの再生成を行う手法について図13を用いて説明する。

ヒント情報解析部61では、規定のフォーマットで記述されたメタデータ最適化ヒント情報60の解析を行う。メタデータ解析・再生成部63では、ヒント情報解析部61より出力される解析されたメタデータ最適化ヒント情報62を使って、メタデータ49の解析を行い、メタデータ再生成に関わる条件65に基づいて再生成したメタデータを出力する。

【0117】

図18に、メタデータ最適化ヒント情報62を使ったメタデータの解析方法の一例を示す。この例では、もとのメタデータ49から重要度が0.5以上の特徴をもつビデオセグメントのみを抽出し、抽出したビデオセグメントに関する記述のみで構成されたメタデータを再生成するものとする。まずメタデータ再生成のための条件65より、再生成に必要なメタデータを特定する(S1)。ここでは、重要度が0.5以上の特徴をもつビデオセグメントのみを抽出するため、「重要度」や「ビデオセグメント」が再生成に必要な記述子である。次にメタデータ最適化ヒント情報62より、メタデータ49にS1で特定した記述子(以下、「重要度」記述子を例に説明する)が含まれるかどうかを判定する(S2)。メタデータに「重要度」の記述子が含まれている場合には、メタデータの解析を行い(S3)、「重要度」の記述子が含まれていない場合には、メタデータの解析処理を終了する(S4)。

【0118】

また、メタデータ最適化ヒント情報62に、「重要度」の出現位置が階層構造のレベル3までと指定されている場合には、レベル3までのビデオセグメントの解析を終了した時点で(S5)、レベル4以下の階層についての解析を行わずに解析処理を終了する(S6)。なお必要に応じて別のメタデータ49の解析を行うため、S1以降の処理を繰り返す。またメタデータ最適化ヒント情報62に「重要度」記述子の出現回数が20と指定されている場合には、20個の「重要度」記述子を解析し終えた時点で(S5)、メタデータの解析を終了する(S6)。なお、S4またはS6でメタデータの解析処理を終了後、必要に応じて別のメタデータの解析を行うため、S1以降の処理を繰り返す。

【0119】

図19に、メタデータ最適化ヒント情報62を使ったメタデータの解析方法の別

の例を示す。この例では、「タイトル」記述子を含むビデオセグメントのみを抽出してメタデータを再生成するものとする。メタデータに「タイトル」記述子を含むかどうかの判定は、図18の例と同じである。メタデータに「タイトル」記述子を含む場合には、メタデータ最適化ヒント情報に記述されている出現位置のIDに一致するビデオセグメントかどうかを判定する(S7)。IDに一致しない場合には、「タイトル」記述子を含まないビデオセグメントのため、そのビデオセグメントの記述の解析をスキップする(S9)。IDに一致する場合には、「タイトル」記述子を取得するため、そのビデオセグメントの記述の解析を行う(S10)。次に出現位置のIDに一致するすべてのビデオセグメントを解析し終えた場合(S9)には、これ以上、「タイトル」記述子を含むビデオセグメントはメタデータ内に存在しないため、解析処理を終了する(S11)。なお、必要に応じて別のメタデータの解析を行うため、S1以降の処理を繰り返す。以上の解析処理を通して抽出された記述子で再構成したメタデータ64を出力する。

【0120】

メタデータ配信部45では、再構成されたメタデータ64を各種クライアント端末へ配信する。

【0121】

なお、図示はしていないが、メタデータ再生後にはメタデータファイルの所在やメタデータファイルサイズ、メタデータに出現する要素数、メタデータ構成要素に関する情報も変わるため、再生成後のメタデータに対応するメタデータ最適化ヒント情報を再生成してもよい。

【0122】

以上のように、従来はメタデータ再生成のためにメタデータに含まれるすべての記述子を解析しなければならなかったが、本実施の形態5においては、メタデータ49に含まれる記述子のリストや記述子の出現位置、出現回数等が記述されたメタデータ最適化ヒント情報60を使ってメタデータ49の記述子を解析を行うようにしたため、メタデータ再生成のためのメタデータ49自体の解析を省くことができ、また出現位置や出現回数により、再生成条件に合致しない記述子の解析を省くことができるため、メタデータの解析及び再生成に伴う処理コスト(処理量や

メモリ使用量等)を低減することができる。

【0123】

実施の形態6.

実施の形態5では、メタデータ再生のためのメタデータ最適化ヒント情報を用いて、メタデータの解析および再生に伴う処理コストを低減する手段について述べたが、本実施の形態6ではメタデータ最適化ヒント情報を用いてメタデータの検索に伴う処理を軽減する手段を備えたメタデータ検索サーバについて述べる。

【0124】

図20に、本実施の形態6におけるメタデータ検索サーバを示す。図20において、60はメタデータ最適化ヒント情報、61はヒント情報解析部、62は解析されたメタデータ最適化ヒント情報、70は検索条件、71はメタデータ解析部、72は解析されたメタデータ、73は検索部、74は検索結果、である。メタデータ最適化ヒント情報60のフォーマット、ヒント情報解析部61は、実施の形態6と同じであるので、説明を省略し、メタデータ解析部71と検索部73の動作を詳述する。

【0125】

本実施の形態6におけるメタデータ解析部71は、解析されたメタデータ最適化ヒント情報62と検索条件70を用いて、コンテンツの構造や特徴を記述した膨大なメタデータ49の解析を少ない処理コストで効率的に行う手段を提供する。

【0126】

図21に、メタデータ解析部71の処理の一例を用いて詳述する。

メタデータ解析部71は、1つ以上のメタデータの解析を各メタデータに対応するメタデータ最適化ヒント情報62を用いて行う。メタデータの解析とは、ここでは、メタデータから検索に必要な特徴記述の抽出を行うことである。例えば、検索条件としてビデオセグメントの色の特徴量が与えられ、そのビデオセグメントに近い特徴をもつビデオセグメントを検索する場合には、色に関する特徴記述をもつビデオセグメントを抽出する必要がある。図15に示すメタデータ例では、レベル4のビデオセグメントに対し、色の特徴記述(「色のヒストグラム」)が付加されているため、レベル4のビデオセグメントに関する記述を抽出する。

【0127】

ここで、図21を参照してメタデータの解析の具体的な処理について詳述する。メタデータ解析部71は、検索条件70を解析し、検索に有効な記述子の特定を行う(S21)。検索条件として、MPEG-7に規定された記述に従った特徴量が与えられる場合と、画像やキーワードが与えられる場合などがある。検索条件がMPEG-7記述に従った特徴量(例えば色配置情報)として与えられる場合には、その記述子(色配置情報)が検索に有効な記述子となる。また、キーワードとして検索条件が与えられる場合には、テキスト形式の記述子(タイトル、アブストラクト、注釈など)が、検索に有効な記述子となる。

【0128】

次にメタデータ最適化ヒント情報62を参照して、選択された記述子がメタデータ49に含まれるかどうかの判定を行う(S22)。検索に用いる記述子がメタデータ49に含まれない場合には(S22 “No”)、そのメタデータ49の解析処理を終了(S24)し、必要に応じて別のメタデータ49の解析を行う。選択された記述子がメタデータ49に含まれる場合には(S22 “Yes”)、メタデータの解析を行う(S23)。メタデータの解析方法については実施の形態5の場合と同様にメタデータ最適化ヒント情報62を用いて図18や図19に示すメタデータ解析処理を効率的に行う(S25～S26)。以上の処理を通して、メタデータ解析部71より検索に必要な特徴記述が抽出される。

【0129】

検索部66では、メタデータ解析部64より出力されるメタデータの解析結果(検索に必要な特徴記述)を利用して、検索条件に適合するコンテンツを検索する。上述した例では、色の特徴記述(「色のヒストグラム」)をもったビデオセグメントに関する記述がメタデータ解析部64より出力されるので、検索条件として与えられた色の特徴量(ヒストグラム)との適合性を判定し、適合したビデオセグメントの情報(例えば「時間情報」)を検索結果として出力する。

【0130】

以上のように、本実施の形態6では、メタデータ最適化ヒント情報60を使ってメタデータ49の解析を行うため、メタデータ再生のためのメタデータ49自体の

解析を省くことができ、また出現位置や出現回数により、検索に必要なではない記述子の解析を省くことができるため、メタデータの検索に伴う処理コスト(処理量やメモリ使用量等)を低減することができる。

【0131】

実施の形態7.

実施の形態6または実施の形態7では、メタデータ最適化ヒント情報62をサーバ側で利用する方法について述べたが、本実施の形態7では、メタデータ最適化ヒント情報をクライアントーサーバ間で利用して、端末やアプリケーションに適切なメタデータを再生成する手法について述べる。

【0132】

図22に、本実施の形態7におけるクライアント端末の構成を示す。なお、図22には、クライアント端末に備わる機能のうち、メタデータ最適化ヒント情報60を利用してメタデータ再生成のための条件を設定する手段に関わる部分のみを示している。図22において、80はヒント情報解析部、81はメタデータ最適化ヒント情報60の解析結果、82はメタデータ再生成条件設定部、83はメタデータ再生のための条件、である。

【0133】

ヒント情報解析部80では、規定のフォーマットで記述されたメタデータ最適化ヒント情報60の解析を行う。

【0134】

メタデータ再生成条件設定部82は、ヒント情報解析部80より出力される解析結果81をもとに、メタデータ再生成の条件設定を行う。条件設定とは、例えば、メタデータ最適化ヒント情報60に含まれる記述子の種類の中からクライアント端末に不必要な記述子の選択がある。クライアント端末が特徴量を用いた検索機能を備えない場合には、色のヒストグラムや動きの複雑性などの特徴量を表現する記述子は不要である。条件設定の別の例として、コンテンツのシーン関係を記述する階層構造が深くなるほどメタデータの複雑度が増すときに、メタデータ最適化ヒント情報60に記述されている階層構造の深さの最大値をもとに、クライアント端末で処理可能な階層構造の深さを設定する。また、別の例では、メタデータ最

適化ヒント情報60に記述されている重要度がとり得る値の情報をもとに、ユーザが着目する視点と、選択するシーンの重要度の閾値を設定する。

【0135】

上述したように、重要度が"Team A"、"Team B"のそれぞれの視点から、5つの離散的な値($\{0.0, 0.25, 0.5, 0.75, 1.0\}$)をとる場合には、"Team A"の視点で0.5以上の重要度をもつシーンのみを選択するなどの設定を行うことができる。メタデータ再生成条件設定部82で設定されたメタデータ再生成のための条件は、メタデータ配信サーバに送信される。サーバ側では、メタデータ再生成のための条件と、クライアント端末の端末性能とに基づいてメタデータを再構成する。例えば、もとのメタデータの階層構造の深さの最大値が4で、メタデータ再生成条件に、クライアント端末で処理可能な階層構造の深さが2と設定されている場合には、階層構造の深さの最大値が2となるように、メタデータの構造を再構成する。また、メタデータ再生成条件に、"Team A"の視点で重要度が0.5以上のシーンのみを選択すると設定されている場合には、同条件に合致するシーンのみで構成されたメタデータを再生成する。メタデータ再生成は、実施の形態5と同様にメタデータ最適化ヒント情報を用いて効率的に行うことができる。

【0136】

以上のように、本実施の形態7では、メタデータ最適化ヒント情報60を使ってメタデータ再生成のための条件を設定することができるため、クライアント端末やアプリケーションに応じて適切なメタデータを生成することができる。

【0137】

実施の形態8.

実施の形態5または実施の形態6では、メタデータ最適化ヒント情報60を用いてメタデータを再生成し、再生成したメタデータを配信する手法について説明したが、本実施の形態8では、メタデータ最適化ヒント情報60を用いてメタデータを解析し、解析結果を利用してクライアント端末やユーザ嗜好に適したコンテンツを再生成して配信するコンテンツ配信サーバについて説明する。

【0138】

図23に、本実施の形態8におけるコンテンツ配信サーバの構成を示す。図2

3において、49はコンテンツの構造や特徴を記述したメタデータ、60はメタデータ最適化ヒント情報、61はヒント情報解析部、62は解析されたメタデータ最適化ヒント情報、85はクライアント端末に関する情報、あるいは、ユーザ嗜好等のコンテンツ再構成に関わる条件、86はメタデータ解析部、87は解析されたメタデータ、88はコンテンツ再構成・配信部、89はもとのコンテンツ、90は再構成されたコンテンツ、である。

【0139】

次に動作を説明する。尚、図23において、ヒント情報解析部61の動作は実施の形態5と同様であるので、説明を省略する。

メタデータ解析部86は、ヒント情報解析部61より出力される解析されたメタデータ最適化ヒント情報62を使って、メタデータ49の解析を行い、コンテンツ再構成に関わる条件に合致する記述を抽出する。ヒント情報を使った解析方法は、実施の形態5と同様であるが、実施の形態5と異なる点は、抽出した記述を使ってメタデータを再生成するのではなく、コンテンツを再構成する点である。メタデータ解析部86で抽出された記述、すなわち解析されたメタデータ87は、コンテンツ再構成部88へ出力される。

【0140】

コンテンツ再構成・配信部88は、メタデータ解析部86で抽出された記述を元に、コンテンツ89の再構成を行う。ここでは、実施の形態5で述べた例を用いて説明する。実施の形態5の例では、メタデータ49から重要度が0.5以上の特徴をもつビデオセグメントのみを抽出し、抽出したビデオセグメントに関する記述のみで構成されたメタデータを再生成したが、本実施の形態8では、メタデータ49から重要度が0.5以上の特徴をもつビデオセグメントのみを抽出し、抽出したビデオセグメントに対応するシーンのみで構成されたコンテンツ90を再構成して配信する。抽出したビデオセグメントに関する記述には、対応するコンテンツ90の所在とそのビデオセグメントのコンテンツ90内での位置(時間情報)が記述されているので、コンテンツ90から対応するシーンを切り出し、一つのコンテンツ90に再構成した上で配信することもできるが、コンテンツ90から対応するシーンを切り出し、切り出したシーンを順次配信することもできる。

【0141】

以上のように、本実施の形態8におけるコンテンツ配信サーバでは、メタデータ49に含まれる記述子のリストや記述子の出現位置、出現回数等が記述されたメタデータ最適化ヒント情報60を使ってメタデータの解析を行うため、メタデータ再生のためのメタデータ49自体の解析を省くことができ、また出現位置や出現回数により、再生条件に合致しない記述子の解析を省くことができるため、クライアント端末やユーザ嗜好に適したコンテンツを再生して配信する際のメタデータの解析及びコンテンツの再構成に伴う処理コスト(処理量やメモリ使用量等)を低減することができる。

【0142】

【発明の効果】

以上説明したように、本発明によれば、動画像や音声を含むマルチメディアコンテンツを複数のシーンに分割し、複数の分割したシーンを編集してマルチメディアコンテンツの階層的な構造を記述するシーン構造情報メタデータを生成するようにしたので、ビデオデータ等を含むマルチメディアコンテンツのもつ階層的な構造を記述したメタデータを生成することができる。

【図面の簡単な説明】

【図1】 実施の形態1におけるメタデータ編集装置の構成を示す図。

【図2】 ニュース映像の一例を示す図。

【図3】 実施の形態1のシーン分割部3が生成したメタデータの一例を示す図。

【図4】 シーン編集部5での編集の結果、生成された映像コンテンツの階層的な構造を記述するメタデータを示す図。

【図5】 コンテンツ再生・表示部2と、ユーザ入力部8の画面イメージ例を示す図。

【図6】 実施の形態2におけるメタデータ再生装置の構成を示す図。

【図7】 実施の形態2における要約されたシーン構造を階層的に記述したものの一例を示す図。

【図8】 実施の形態3におけるメタデータ編集装置の構成を示す図。

【図9】 実施の形態3においてシーンチェンジ点を自動検出する一例を示す図。

【図10】 実施の形態4におけるコンテンツ配信システムの構成を示す図。

【図11】 ニュース映像を例に、メタデータ解析部42より出力されるコンテンツの構造情報の一例を示す図。

【図12】 実施の形態4による再構成後のコンテンツの構造例を示す図。

【図13】 実施の形態5におけるメタデータ配信サーバの構成を示す図。

【図14】 メタデータ最適化ヒント情報について詳述するためのビデオコンテンツの一例を示す図。

【図15】 MPEG-7を用いた場合のメタデータの記述例を示す図。

【図16】 メタデータ最適化ヒント情報のフォーマット例を示す図。

【図17】 図16に示すフォーマットに従って記述されたメタデータ最適化ヒント情報を示す図。

【図18】 ヒント情報を使った解析方法の一例を示す図。

【図19】 ヒント情報を使った解析方法の他の例を示す図。

【図20】 実施の形態6におけるメタデータ検索サーバを示す図。

【図21】 メタデータ解析部71の処理の一例を用いて詳述する図。

【図22】 実施の形態7におけるクライアント端末の構成を示す図。

【図23】 実施の形態8におけるコンテンツ配信サーバの構成を示す図。

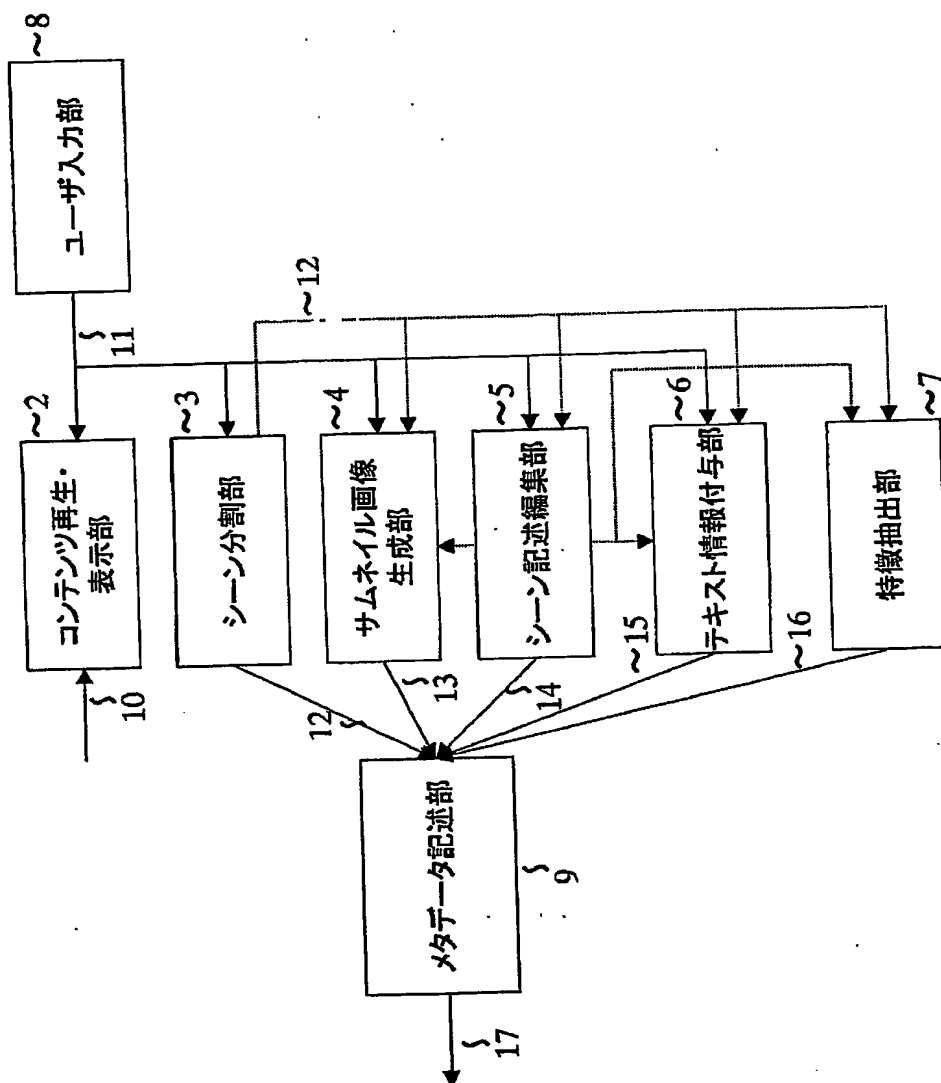
【符号の説明】

2 コンテンツ再生・表示部、3 シーン分割部、4 サムネイル画像生成部、
5 シーン編集部、6 テキスト情報付与部、7 特徴抽出部、8 ユーザ入力部。

【書類名】

図面

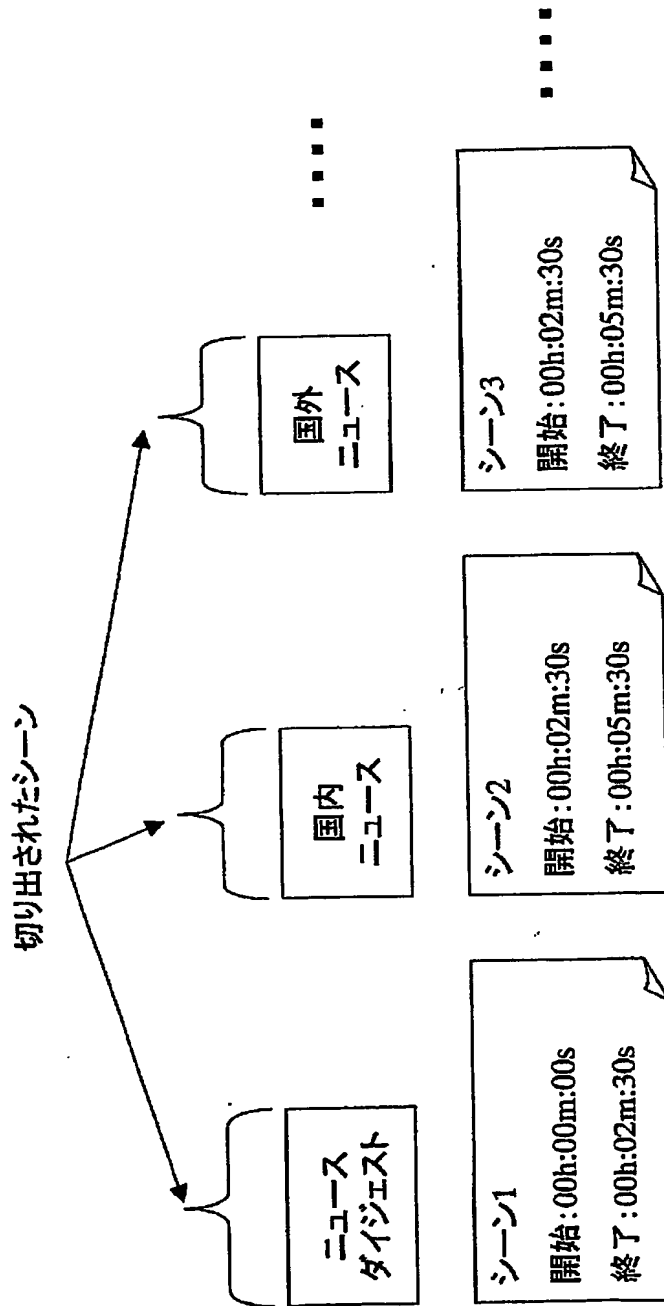
【図 1】



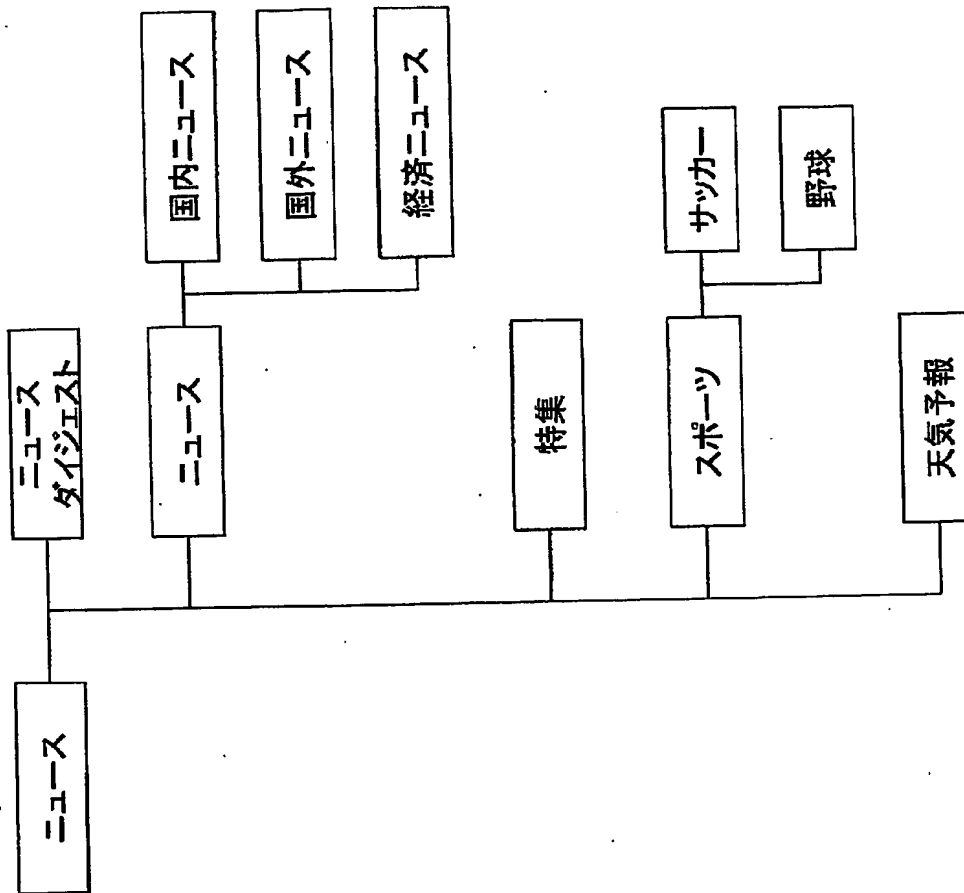
【図 2】

ニュース ダイジェスト	国内 ニュース	国際 ニュース	経済 ニュース	特集	スポーツ サッカー	スポーツ 野球	天気予報
----------------	------------	------------	------------	----	--------------	------------	------

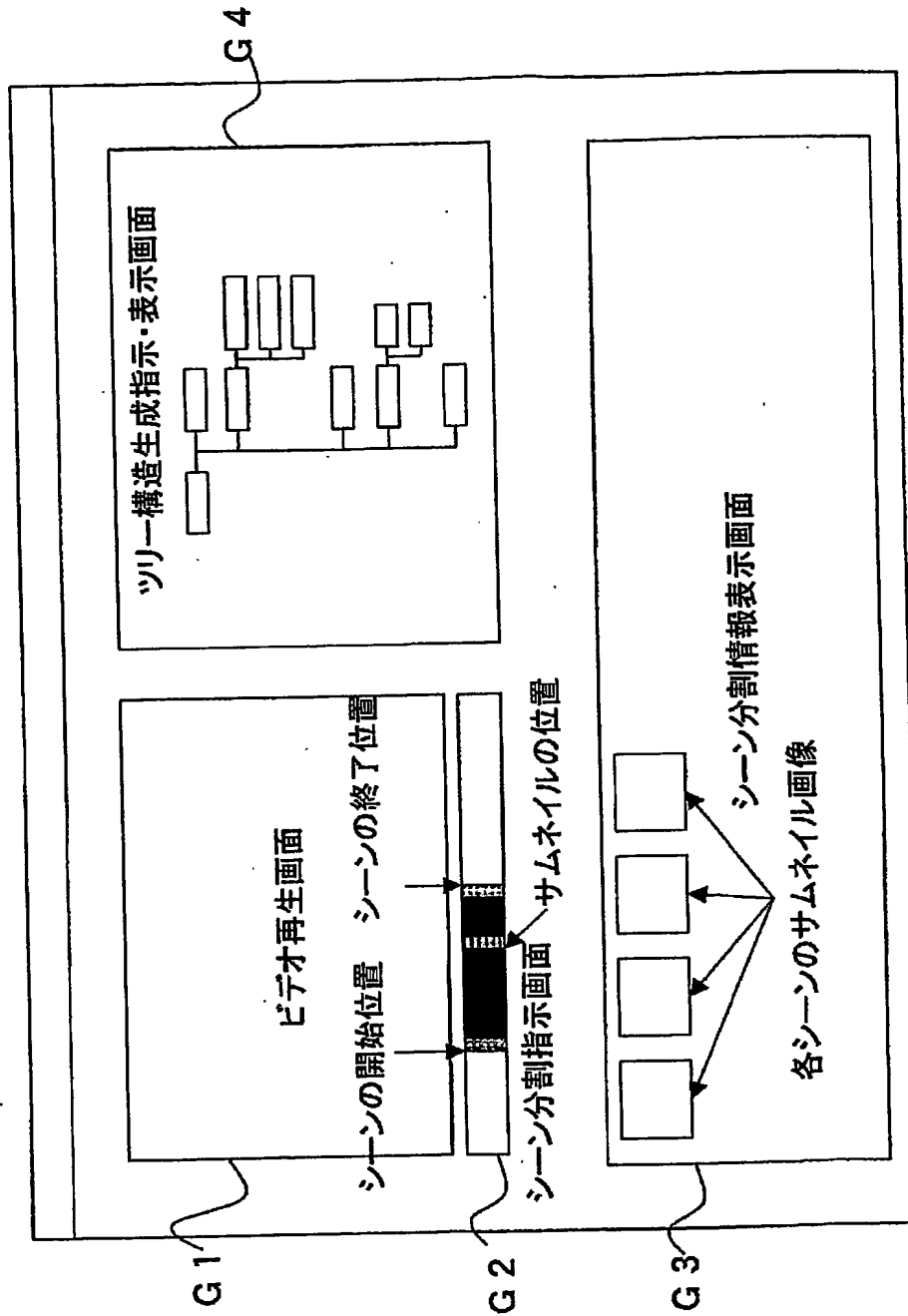
【図3】



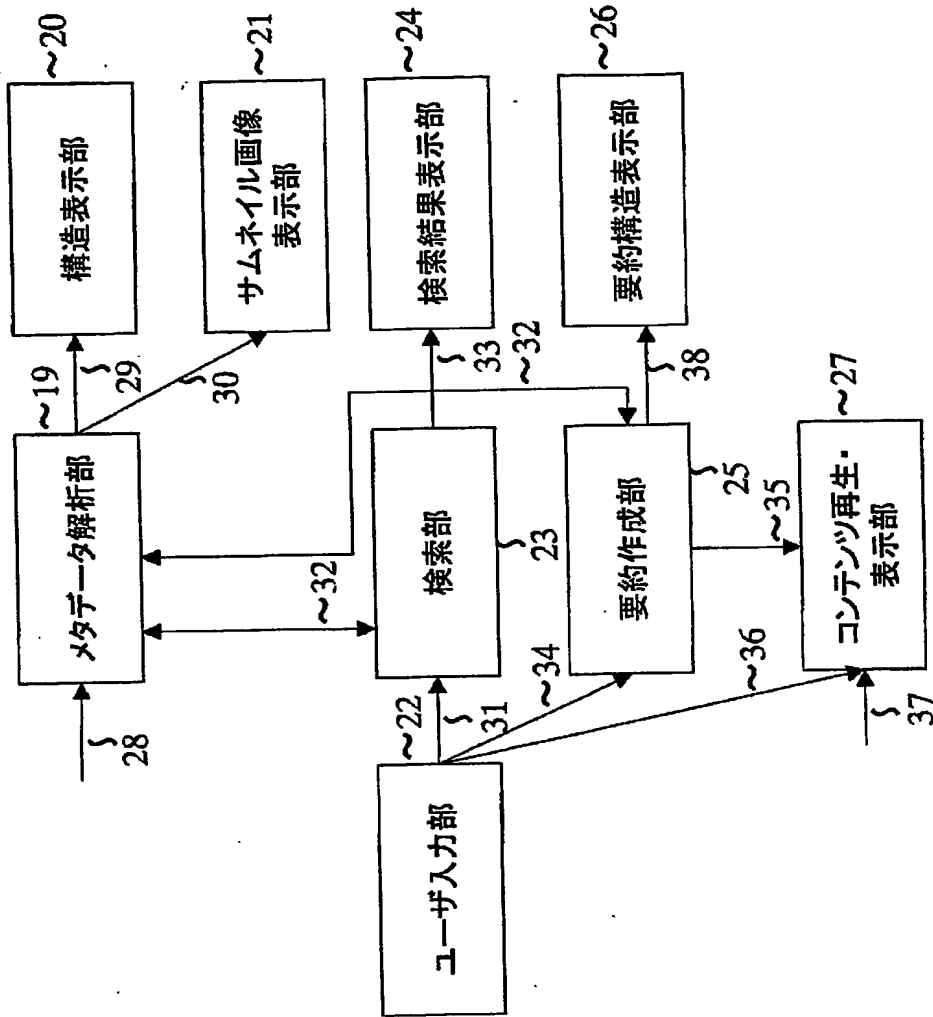
【図4】



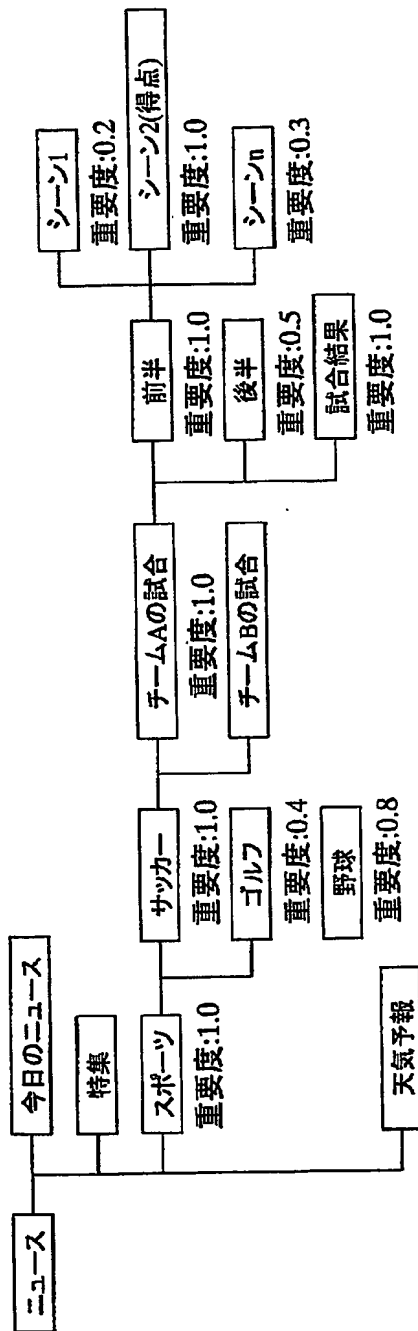
【図 5】



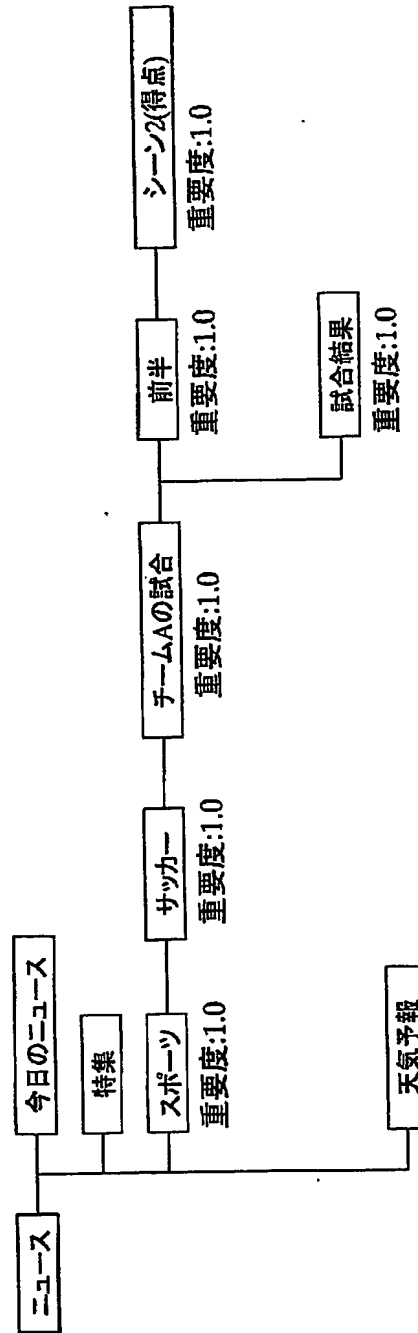
【図 6】



【図 7】

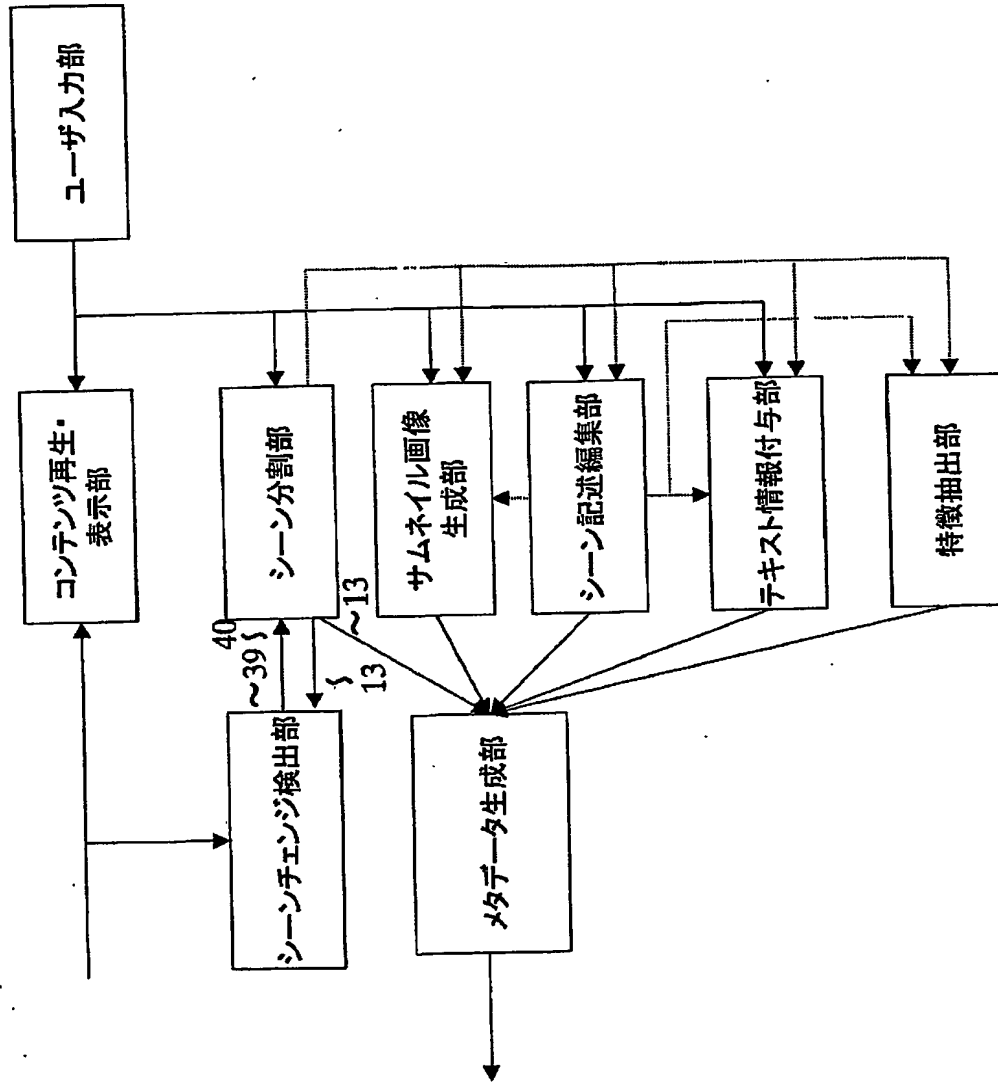


(a)オリジナルコンテンツのシーン構造

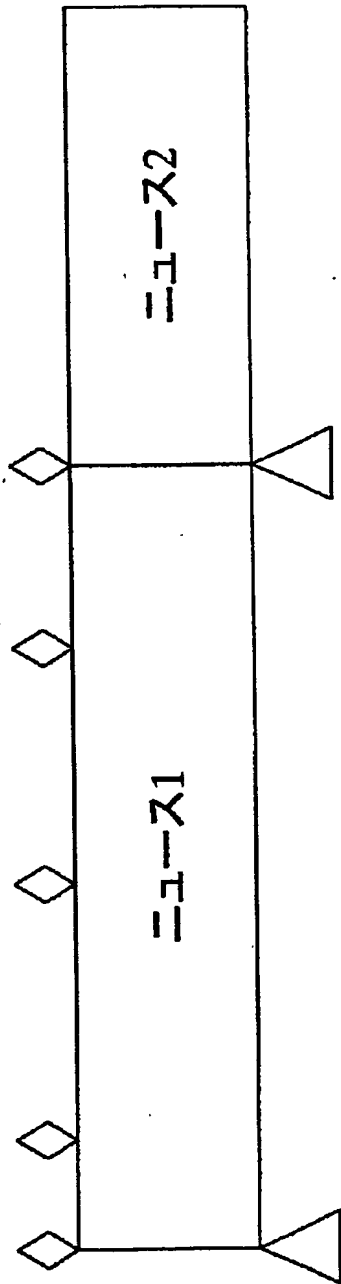


(b)要約されたコンテンツのシーン構造

【図 8】



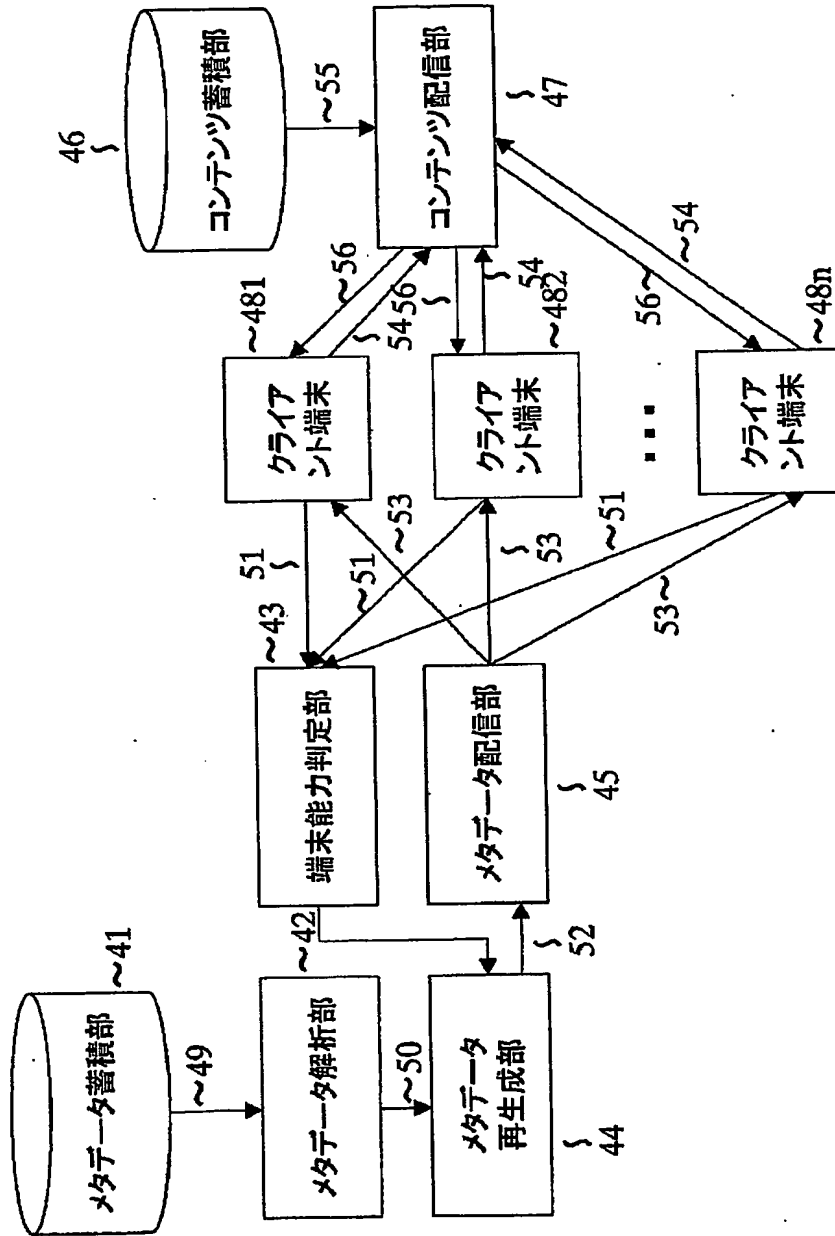
【図9】



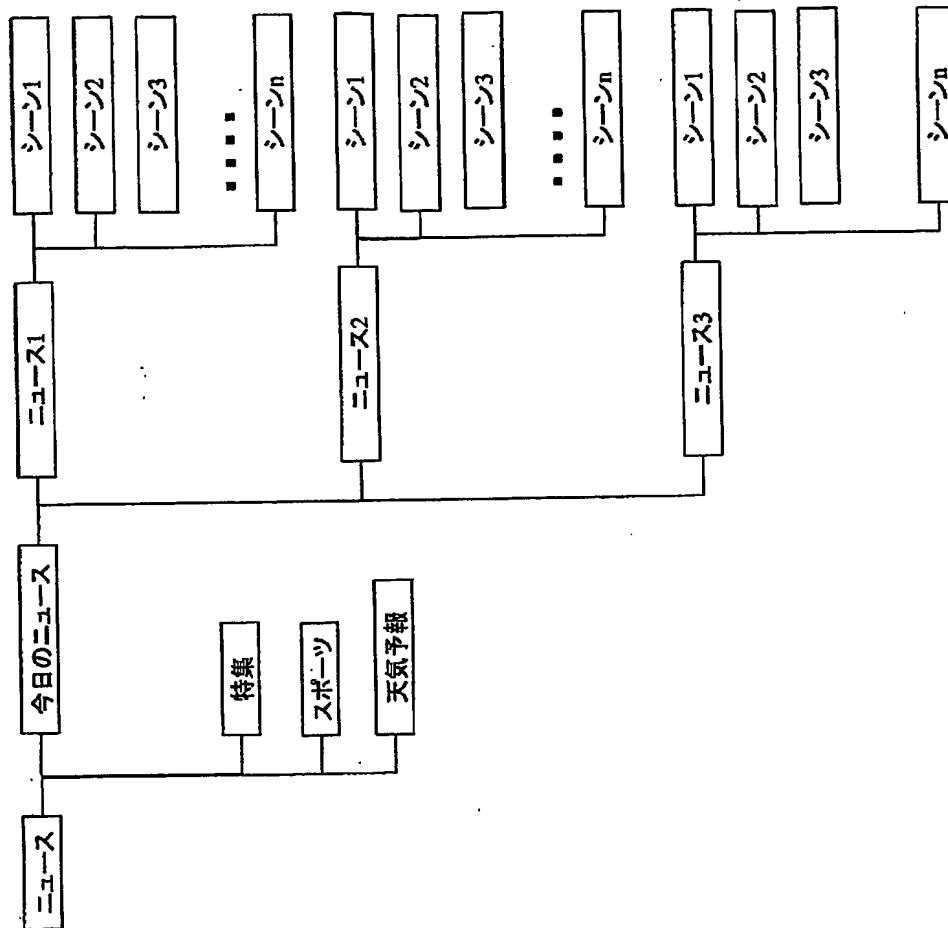
◇ シーンチェンジ点

△ シーン開始点(テンプレート画像とマッチングする点)

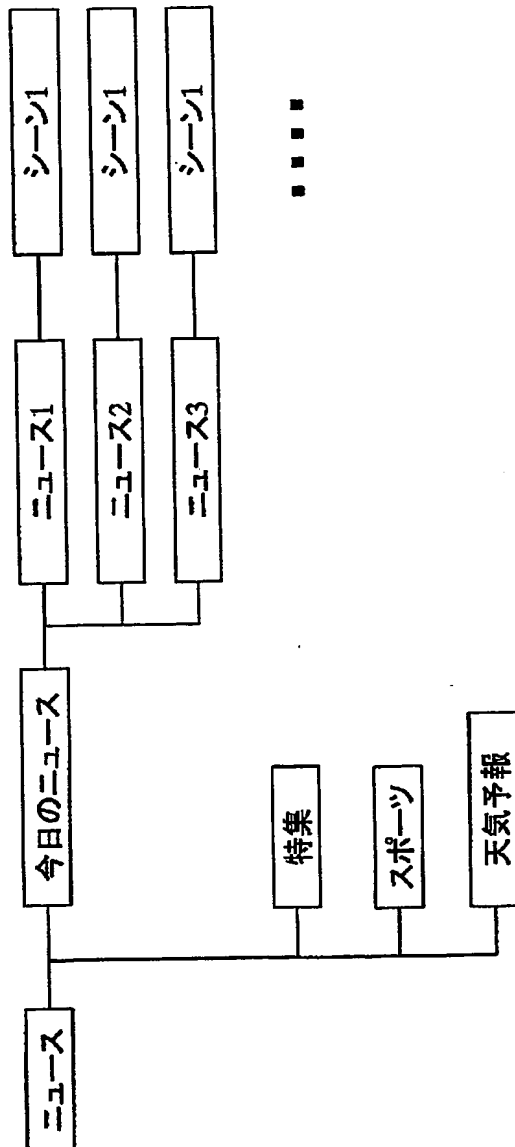
【図10】



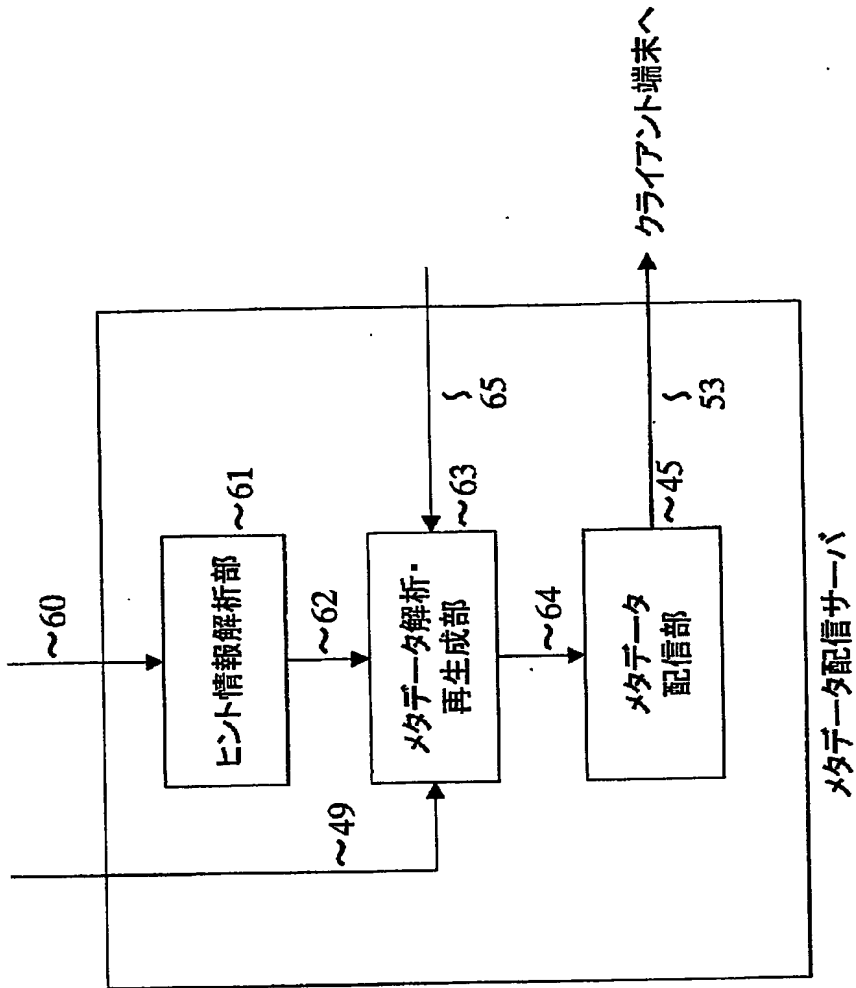
【図 11】



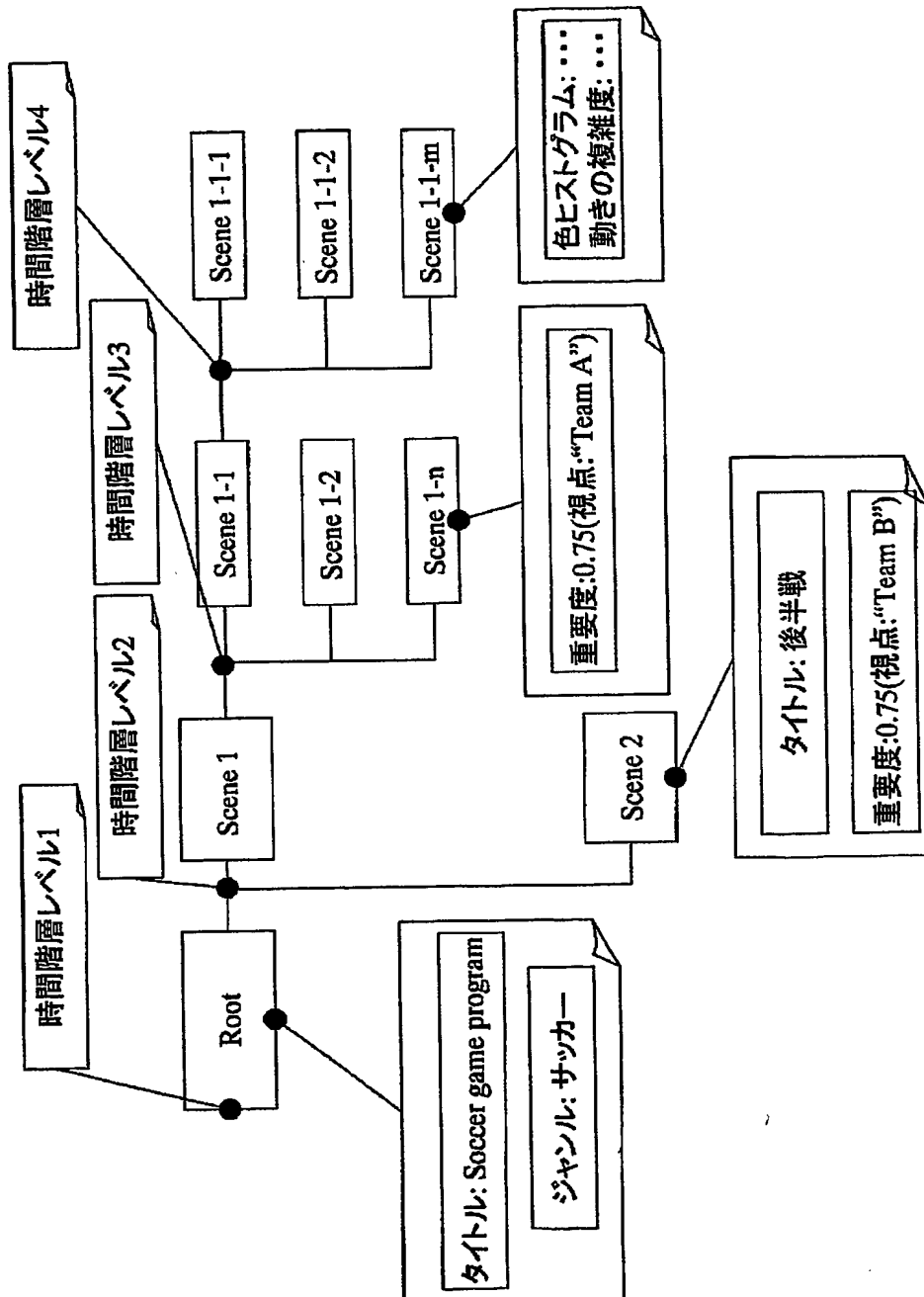
【図12】



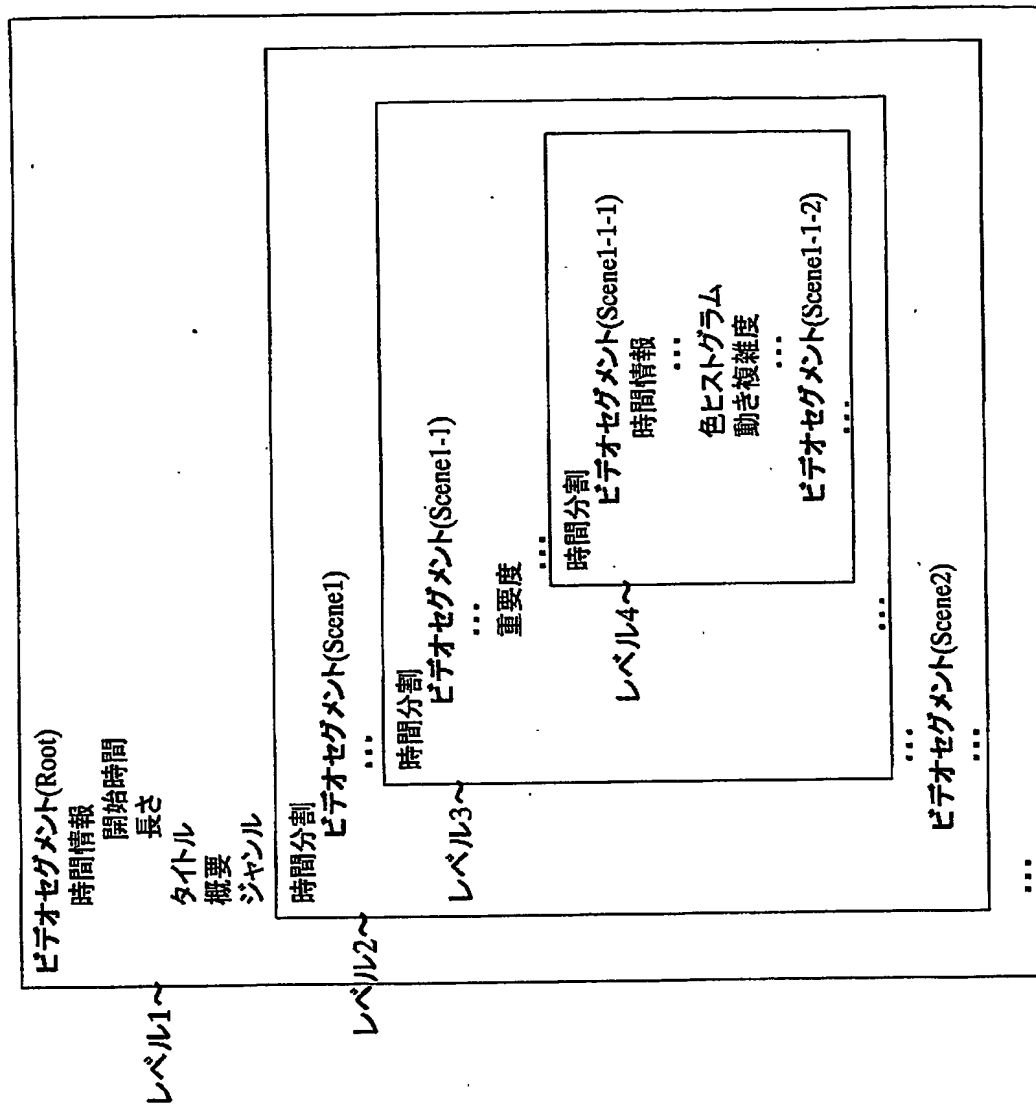
【図 13】



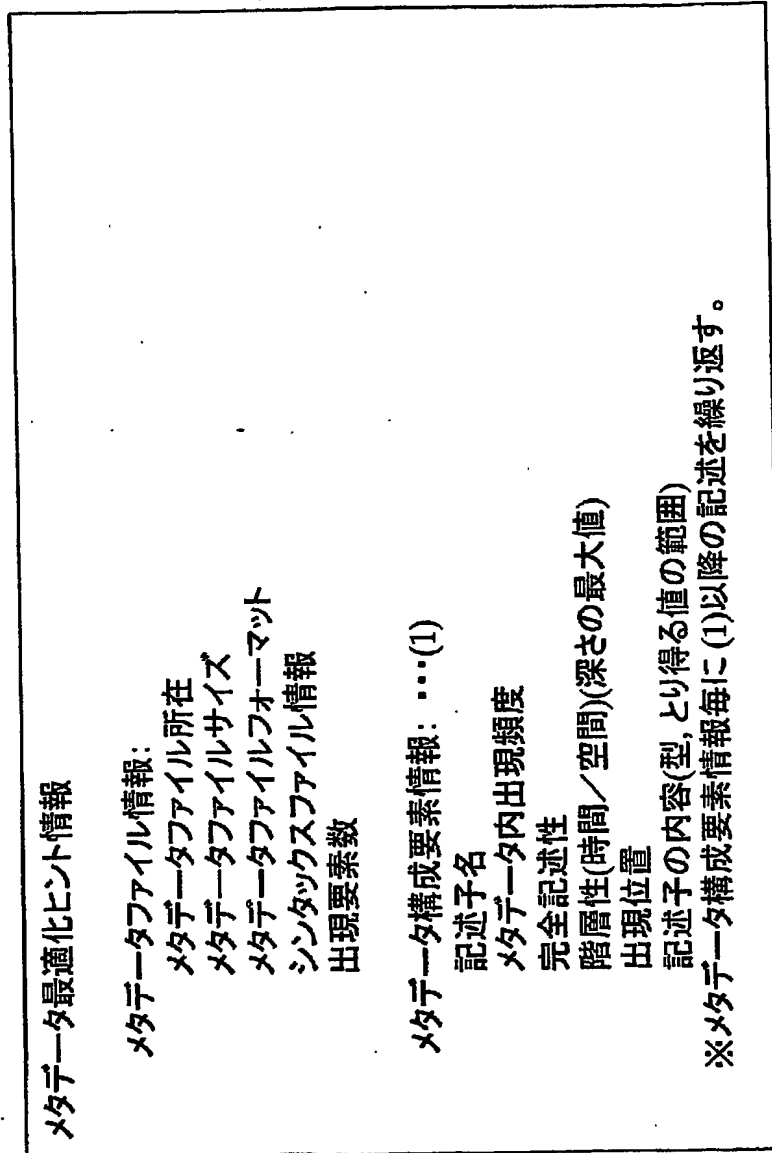
【図14】



【図 15】



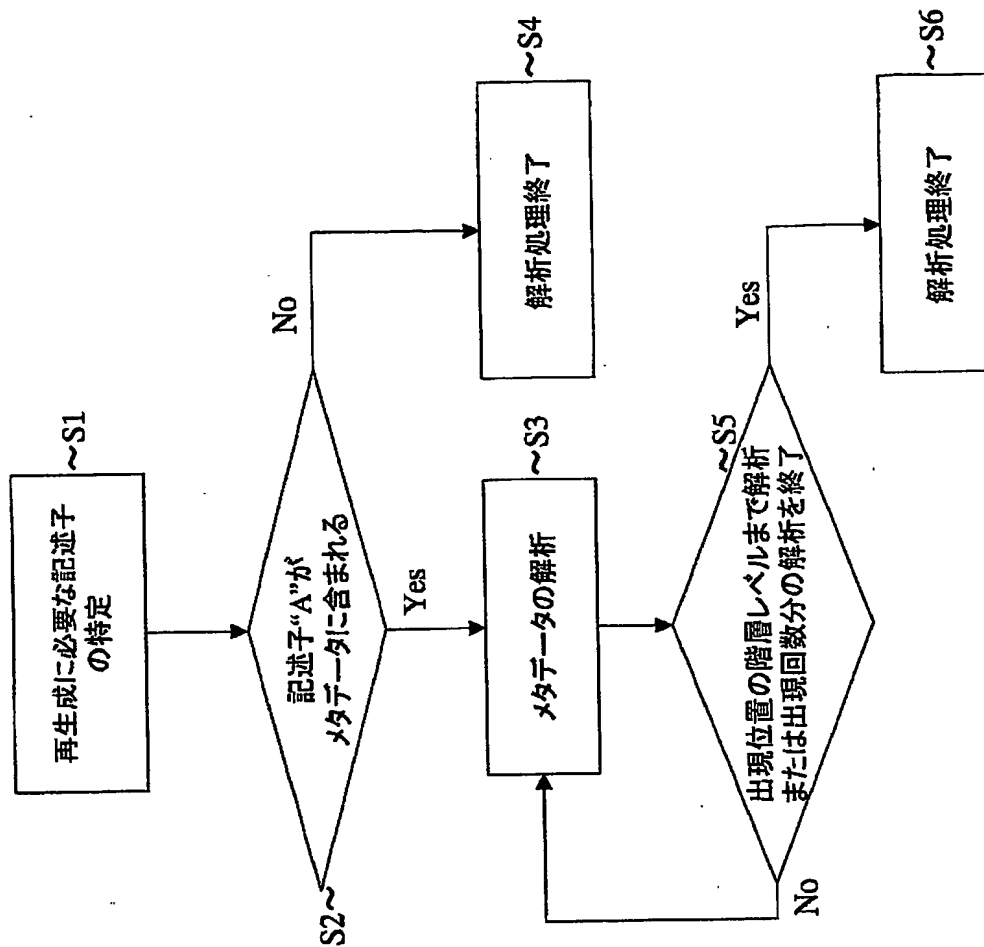
【図 16】



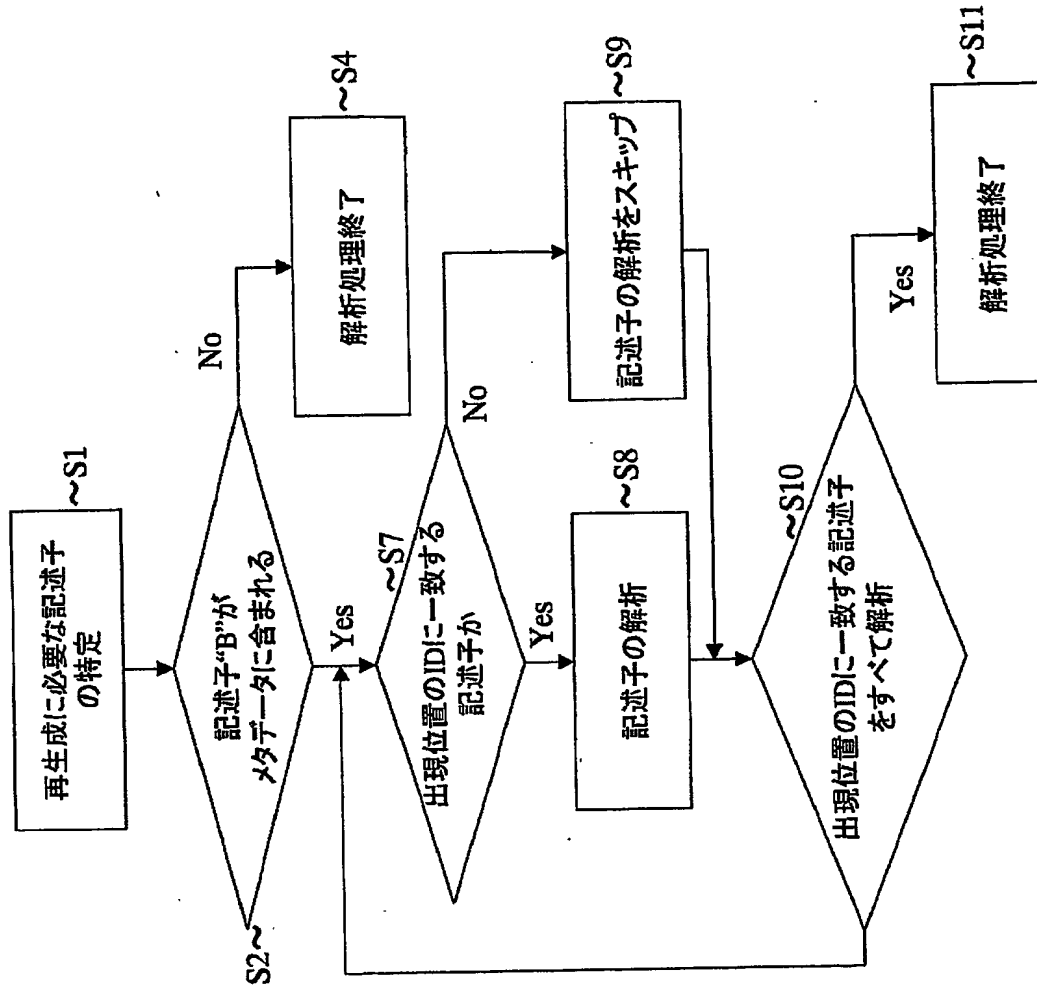
【図 17】

メタデータ最適化ヒント情報
 メタデータファイル情報:
 メタデータファイル所在: `http://xxx.xxx.xx.xx/xxx/xxx`
 メタデータファイルサイズ: 120KB
 メタデータフォーマット: XML
 シンタックスファイル情報: `urn:mpeg:mpeg7:schema:2001 Mpeg7-2001.xsd`
 出現要素数: 150
 メタデータ構成要素情報:
 記述子名: ビデオセグメント
 メタデータ内出現頻度: 30
 完全記述性: No
 時間階層性(深さの最大値): 4
 空間階層性(深さの最大値): 1
 メタデータ構成要素:
 記述子名: タイトル
 メタデータ内出現頻度: 3
 出現位置(IDリスト): "Root", "Scene1", "Scene2"
 記述子の内容:
 データ型: 文字列型データのリスト
 値: "Soccer game program", "前半戦", "後半戦"
 メタデータ構成要素情報:
 記述子名: 重要度
 メタデータ内出現頻度: 20
 完全記述性: No
 出現位置(時間階層レベル): 3
 記述子の内容:
 データ型: 浮動小数点型データのリスト
 値: 0.0, 0.25, 0.5, 0.75, 1.0

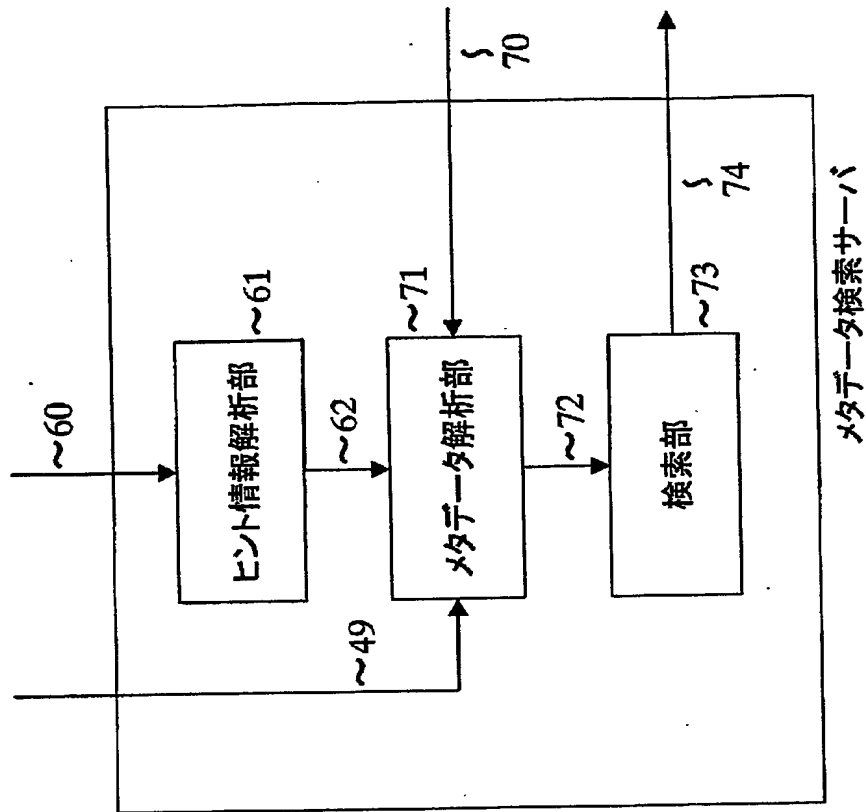
【図18】



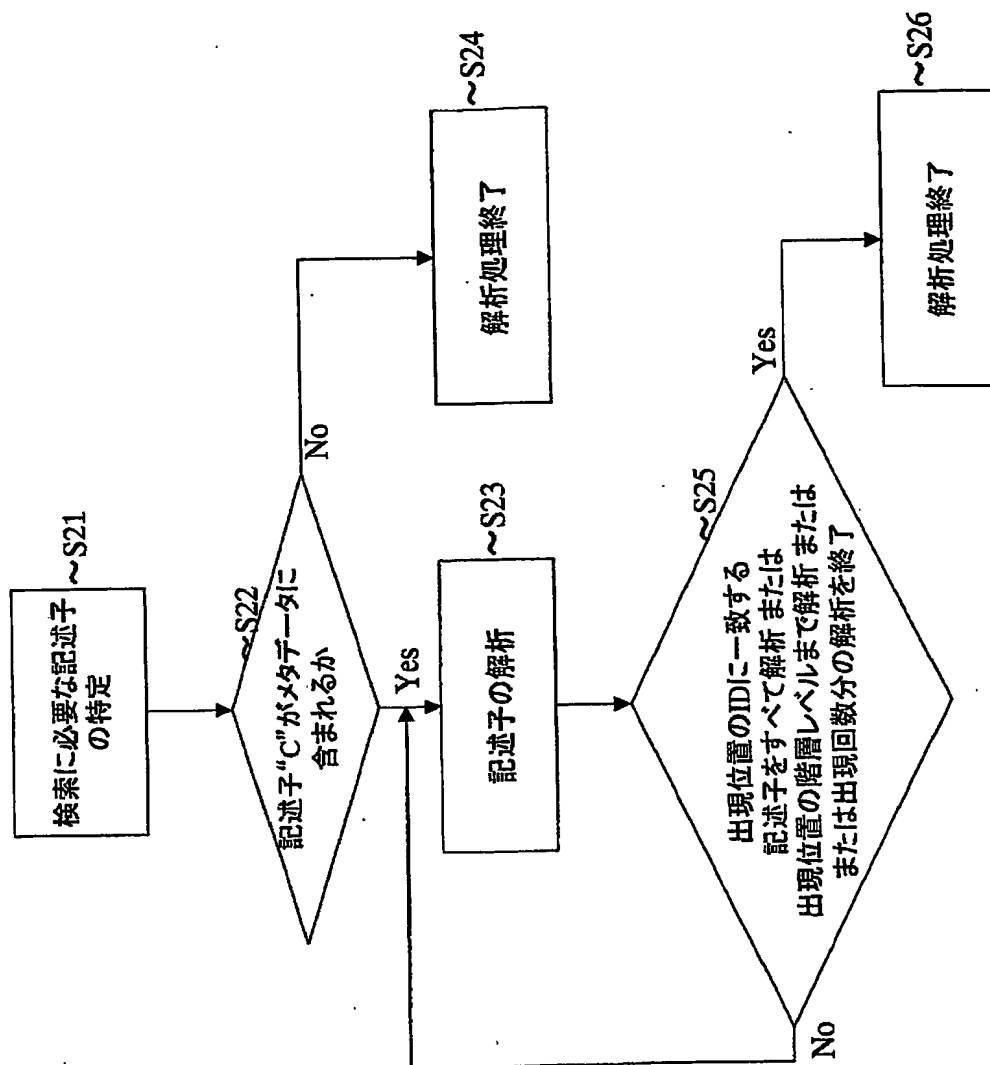
【図 19】



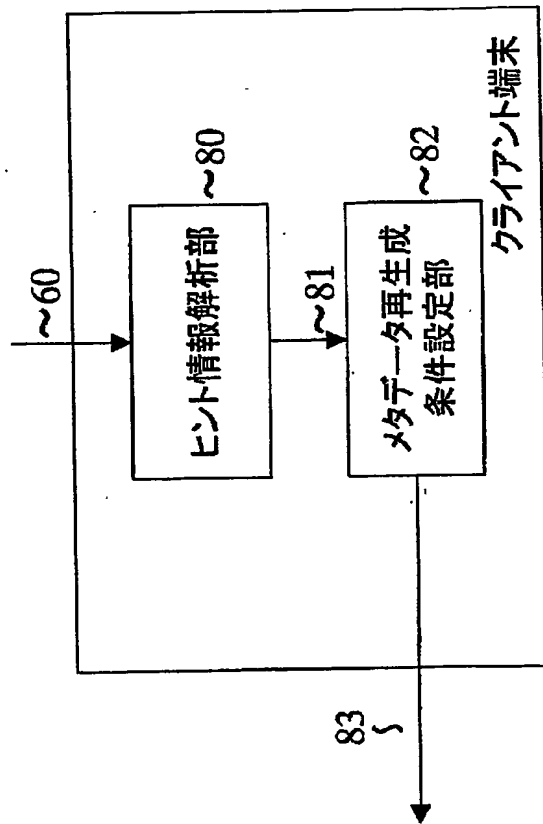
【図 20】



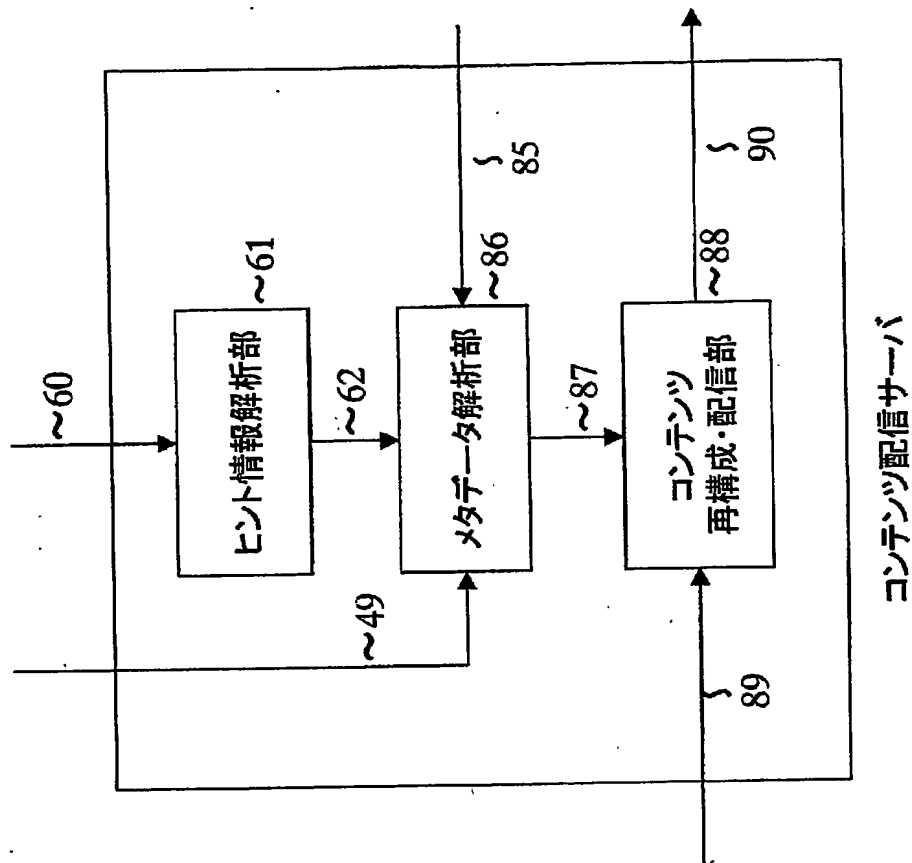
【図 21】



【図 2 2】



【図 23】



【書類名】 要約書

【要約】

【課題】 シーンの区間情報やタイトルのほかに、ビデオデータ等のコンテンツのもつ構造等を含むインデックス情報であるメタデータを生成する。

【解決手段】 映像コンテンツ等のマルチメディアコンテンツ10がネットワーク等を介し本装置に入力すると、コンテンツ再生・表示部2にて再生・表示され、シーン分割部3がユーザ入力部22から入力されたシーンの開始位置と終了位置とを示すシーンの区間情報メタデータ12を生成する。シーン編集部5はユーザ入力部8を介しシーン編集の指示があると、シーン分割部3からのシーンの区間情報メタデータ12に基づき、シーン分割部3で連続的に切り出されたシーンの階層的な編集を行ないシーン構造情報メタデータ14としてメタデータ記述部9へ出力し、メタデータ記述部9は、各種メタデータを統合し、MPEG-7規定の記述フォーマットに従って記述したメタデータファイルを生成する。

【選択図】 図1

認定・付加情報

特許出願の番号	特願2002-178169
受付番号	50200889635
書類名	特許願
担当官	伊藤 雅美 2132
作成日	平成14年 6月26日

<認定情報・付加情報>

【特許出願人】

【識別番号】	000006013
【住所又は居所】	東京都千代田区丸の内二丁目2番3号
【氏名又は名称】	三菱電機株式会社

【代理人】

申請人

【識別番号】	100102439
【住所又は居所】	東京都千代田区丸の内2丁目2番3号 三菱電機株式会社内

【氏名又は名称】	宮田 金雄
----------	-------

【選任した代理人】

【識別番号】	100092462
【住所又は居所】	東京都千代田区丸の内2丁目2番3号 三菱電機株式会社内

【氏名又は名称】	高瀬 彌平
----------	-------

出 願 人 履 歴 情 報

識別番号 [000006013]

1. 変更年月日	1990年 8月24日
[変更理由]	新規登録
住 所	東京都千代田区丸の内2丁目2番3号
氏 名	三菱電機株式会社